Improving Learning Through Assessment Rubrics:

Student Awareness of What and How They Learn

Chahna Gonsalves
King's Business School, King's College London, UK

Jayne Pearson
King's Academy, King's College London, UK

IGI Global
PUBLISHER of TIMELY KNOWLEDGE

Advances in Educational Marketing, Administration, and Leadership (AEMAL) Book Series

Siran Mukerji
IGNOU, India
Purnendu Tripathi
IGNOU, India

MISSION

With more educational institutions entering into public, higher, and professional education, the educational environment has grown increasingly competitive. With this increase in competitiveness has come the need for a greater focus on leadership within the institutions, on administrative handling of educational matters, and on the marketing of the services offered.

The **Advances in Educational Marketing, Administration, & Leadership (AEMAL) Book Series** strives to provide publications that address all these areas and present trending, current research to assist professionals, administrators, and others involved in the education sector in making their decisions.

COVERAGE

- Educational Finance
- Marketing Theories within Education
- Direct marketing of educational programs
- Educational Marketing Campaigns
- Governance in P-12 and Higher Education
- Academic Pricing
- Students as Consumers
- Educational Management
- Educational Leadership
- Faculty Administration and Management

IGI Global is currently accepting manuscripts for publication within this series. To submit a proposal for a volume in this series, please contact our Acquisition Editors at Acquisitions@igi-global.com or visit: http://www.igi-global.com/publish/.

Titles in this Series

Table of Contents

Section 1
Using Rubrics in Context

Section 2
Developing Rubrics

Section 3
New Perspectives on Rubrics

Detailed Table of Contents

Section 1
Using Rubrics in Context

An overview of how rubrics are used- the different types employed for different assessment settings. This covers a range of disciplines, HE and schools, and different countries. This section contains much of the introductory descriptive stuff about rubrics and is best placed first.

Chapter 1

    Laura E. Mendoza, University of Texas at El Paso, USA

The use of rubrics for classroom assessment has been evolving rapidly during the last decades. Unfortunately, because rubrics across language classrooms are only sometimes standardized, some may provide a non-friendly language for students, which is useless for their linguistic development. In addition to defining rubrics from an array of perspectives, the present chapter presents possible benefits linked to the appropriate use of rubrics among emergent bilinguals. Shabani and Panahi highlight essential benefits for students when presented with language assessment tools, such as rubrics, and the authors use their observation for this chapter. It aims to highlight the importance of rubrics in the language classroom in an effort for educators, evaluators, and policymakers to be more conscious when creating and using rubrics.

Chapter 2

    Işıl Aykutlu, Hacettepe University, Turkey

In physics classes, graphs are utilized for various topics to make sure students better understand the relationship between variables. Graphs are especially useful for showing the relationship between dependent and independent variables in experiments carried out in lab classes. There are various sub-steps to drawing a graph. In order to draw a graph, students should have a good grasp of physics subjects; moreover, they should know the characteristics of graphs and be able to form the graph according to the steps. Using rubrics to assess students' drawings at the end of this process, which involves multiple skills, would ensure that the assessment is both more detailed and more reliable and valid. Moreover, using rubrics in teaching how to draw a graph would contribute to students' understanding of the graphs' characteristics; it would also help students clearly see the steps as well as contribute to the improvement

of students' drawings through self and peer assessment.

Chapter 3
Minna Hämäläinen, LUT University, Finland
Anu Raappana, LUT University, Finland

This chapter describes a tool for assessing the development of learners' entrepreneurship competence and its construction process. Competence in entrepreneurship is promoted at all school levels, but there are very few tools for teachers to assess the level or development of students' entrepreneurship competence. It can also be very unclear to teachers what should be assessed and when to assess a student's entrepreneurship competence. The rubric-based tool allows teachers to assess the development of a learner's entrepreneurship competence in an easy and simple way. The rubric also makes it possible to develop a teacher's own skills, as it gives an idea of what is being assessed when it comes to entrepreneurship competence. The tool was developed in a participatory process with teachers, students, and researchers. Another special feature of the presented tool is its internationality and the common contribution of different cultures to the development of the assessment tool.

Chapter 4
Emma O. Brien, Department of Language and Literacy Education, Faculty of Education,
Mary Immaculate College, Limerick, Ireland
Josephine Brady, Department of Language and Literacy Education, Faculty of Education,
Mary Immaculate College, Limerick, Ireland
T. J. Ó Ceallaigh, School of Education, University College Cork, Cork, Ireland
Katharine Babbitt, Department of Language and Literacy Education, Faculty of Education,
Mary Immaculate College, Limerick, Ireland
Andrea Brosnan, Department of Language and Literacy Education, Faculty of Education,
Mary Immaculate College, Limerick, Ireland
Emma Byrne, Department of Language and Literacy Education, Faculty of Education, Mary
Immaculate College, Limerick, Ireland
Erin Byrne, Department of Language and Literacy Education, Faculty of Education, Mary
Immaculate College, Limerick, Ireland
Rebecca Curtin, Department of Language and Literacy Education, Faculty of Education,
Mary Immaculate College, Limerick, Ireland
Lisa Gaffney, Department of Language and Literacy Education, Faculty of Education, Mary
Immaculate College, Limerick, Ireland
Karen O'Callaghan, Department of Language and Literacy Education, Faculty of
Education, Mary Immaculate College, Limerick, Ireland

Developing the teacher as a reflective practitioner has become a core facet of Irish teacher education, but large gap exists between theory and practice. Research illustrates the positive value of rubrics in terms of student self-reflection and self-regulation. However, few studies explore the use of assessment rubrics within the broad context of the supervisory relationship. Drawing on Drytons extended supervisory working alliance the authors explored how rubrics can foster student-supervisor relationships during the

research process. This study adopted a collaborative autoethnographic (CAE) methodology which enabled the researchers to authentically capture the student and faculty perspectives. It was found that rubrics provided transparency, identified expectations and a language for students to express and interrogate their work. They provided a metadialogue to enable students to take ownership of the feedback process questioning and initiating discussion with the supervisor. This supported the bonding process, shaping conversations and providing the student with context for the feedback.

Daniel Bores-García, Rey Juan Carlos University, Spain
Raúl A. Barba-Martín, University of León, Spain
Gustavo González-Calvo, University of Valladolid, Spain
David Hortigüela-Alcalá, University of Burgos, Spain

This chapter presents the analysis of a formative and shared assessment experience in the subject of physical education in secondary education in a high school in Spain. In the assessment process, an assessment scale has been used as an assessment instrument, by means of which students have self-assessed themselves, have been co-assessed by their peers and have received a hetero- assessment from the teacher based on the criteria previously established in the instrument. After the implementation of this experience, a study was carried out on the students' perception of the formative and shared assessment process and the use of the evaluation scale. A discussion group was held, and the teachers' and students' diaries were analyzed. The results show a feeling of motivation and commitment in the students to the task and the group thanks to the feeling of being part of the process through the assessment.

Elena Ramona Richiteanu-Nastase, Bucharest University of Economic Studies, Romania
Alexandru Robert Mihaila, Bucharest University of Economics Studies, Romania

This chapter clarifies first a number of concepts such as evaluation, traditional evaluation methods, alternative evaluation methods, process-centered evaluation, and evaluation of student progress and portfolios as an alternative evaluation method. The authors will approach the concept of rubrics as a very useful evaluation tool, highlighting design and exemplification ways for some more commonly used evaluation methods. Their advantages and disadvantages will be analyzed. After taking into consideration the most important issues and controversies, the authors will analyze an example of good practice, namely the use of rubrics in the evaluation of students with the help of e-portfolios, as an alternative method of evaluation. The last part of the chapter is dedicated to discussions and recommendations for using rubrics in evaluation. Aspects regarding the usefulness, but also the limits of the instrument, as well as ways of further development, will be discussed.

Section 2
Developing Rubrics

The process of development with colleagues and with students. How this can be done and guides for collaborative process and evaluation in different contexts?

Chapter 7

Allan Stephen Laville, University of Reading, UK
Lindsey Thompson, University of Reading, UK
Yue Yue, University of Reading, UK
Alexandra J. Hayward, University of Reading, UK
Victoria Grace-Bland, University of Reading, UK

The chapter explores the importance of utilising student-staff partnerships in the development and evaluation of rubrics. The approach followed is underpinned by the University of Reading principles for student-staff partnerships that centres student voice in the development of Teaching and Learning initiatives. The chapter explores the challenges of engaging students with assessment rubrics and through engaging in listening exercises, the actions taken to remove these barriers to engagement. The chapter provides three case studies that detail practical recommendations to improving student assessment literacy including in-class support for rubrics, additional support outside of the classroom e.g., assessment rubric screencasts and discussion boards, and the importance of co-creation in creating new rubrics. The chapter concludes by detailing the importance of student-staff partnerships in rubric development and evaluation, but also detailing the additional support mechanisms that need to be in place to effectively develop student and staff assessment literacy.

Chapter 8

Anja Harrison, King's College London, UK
Maren Breier, King's College London, UK
Harriet Power, King's College London, UK
Brenda P. Williams, King's College London, UK

Marking rubrics is hailed as a transparent and effective way of supporting student success; enhancing their ability to understand and use their learning environment to achieve their goals. Rubrics also enable staff to mark fairly and consistently. Yet, to be successful, rubrics must be understandable to all and there needs to be active engagement from students and staff alike. Understandability requires that the wording be inclusive and considerate of student diversity, and this is especially true when considering online courses where the student body is often more culturally diverse. Co-creation with students can promote inclusivity and the development of meaningful and successful rubrics. This chapter provides a step-by-step guide for co-creating and implementing rubrics in an online education context, developed through collaboration with the co-creation student panel from the online programmes at the Institute of Psychiatry, Psychology and Neuroscience, KCL.

     Roisin Donnelly, Technological University Dublin, Ireland
     Colin Hughes, Technological University Dublin, Ireland

This chapter is a reflective study reporting on a College-wide common rubrics initiative in a Technological University (TU) in Ireland. Assessment and feedback are enduring issues for the higher education sector both in Ireland (as well as internationally). By addressing these priorities, we are focusing on the connected areas of marking practices and feedback processes in a College of Business. The chapter highlights the collaborative nature of an initiative on programmatic assessment design, its breadth of scope, and the high levels of support provided to staff and students through the design process. In particular, rubrics are the main focus of the chapter with an overview of Programme Learning Outcome (PLO) mapping provided as part of the context. Four interdisciplinary rubric working teams were formed across the College of Business to develop common rubrics in the areas of reflective practice, critical thinking, individual/ group presentation skills and industry consultancy projects. This programme-based study differs from other previous work as it involves working on a consistent basis with the challenges of bringing cultures, practices and understandings of disciplinary teams together in a technological university context. Findings from our collaborative common rubric working teams showed the importance of avoiding designing overly complex rubrics and of focusing on the student's work providing evidence of meeting, exceeding, or falling short of the quality being looked for. It also highlighted the importance of rubric literacy as we were using qualitative language in all the target statements and advised staff to avoid subjective language in favour of explicit guidance. We proposed the need for a common language on 'Teamwork' and on proceeding with this as a College-wide approachd a similar structure or focus, using the same marking rubric was recommended. We advised staff to work as a team to master the rubrics, and to ensure that their students had m. Future work can explore the use of AI tools which can automate the feedback process and provide lecturers with customised rubrics based on their specifications.

     Ömer Açıkgöz, Social Sciences University of Ankara, Turkey
     Aydın Aslan, Selcuk University, Turkey
     Korkut Koçak, Republic of Türkiye Ministry of National Education, Turkey
     Aslı Günay, Social Sciences University of Ankara, Turkey
     Nevzat Yavuz, The Scientific and Technological Research Council of Türkiye, Turkey

This study aims to develop a rubric to evaluate the dissertations implemented in the fields of educational and social sciences. In the development of this rubric, the acquisition requirements concerning knowledge, skills, and competence at the doctorate level in the European and Turkish qualifications frameworks, the legal framework of Turkish higher education, and the perceptions of 12 experts in the fields of educational and social sciences concerning the common competences of the dissertations were considered. The rubric can contribute to the evaluation of dissertations completed in the field of educational and social sciences concerning these dimensions and provide PhD students, researchers, and academics with a guide to evaluate their academic studies based on an empirical instrument.

Section 3
New Perspectives on Rubrics

Moving the field forward through new perspectives, emerging contexts. Comments on limitations and potential of rubrics and points to ways forward.

Chapter 11

Rubrics have become popular tools for assessment and instruction in higher education. However, language choice and rubric efficacy is a topic that has been largely overlooked in the literature and teacher professional development. Composing an effective rubric—particularly for instructional and formative contexts—is a complex task that requires teachers to think about the implications of their linguistic choices for students' awareness of what and how they learn. In this chapter, the author offers a review of current research and guidance on effective rubric language. Second, this chapter uses the theory of systemic functional linguistics (SFL) to explain how SFL-informed training in rubric design can foreground language considerations to enhance teachers' capacities in effective rubric design. Overall, this chapter demonstrates that developing teachers' knowledge about language and in turn their academic and assessment literacy, is key to developing both types of literacy in students.

Chapter 12

This chapter focuses on self-assessment as a decisive assessment for students' learning and offers the results of a study done with 72 university students. The authors investigated preservice teachers' views on self-assessment, the instruments used, and the implications for their students' learning. The results showed that while preservice teachers appear to understand what self-assessment is theoretically, in practice it is a non-systematic assessment with a planned day and criteria that is invariably quantitative. It appears that it is done because it is legally required and not because teachers believe students need to critically reflect on their learning.

Chapter 13

This chapter outlines an approach to using rubrics in a feedforward process using dialogue between teachers and students that takes place before students hand in their assignments. As such, it aims to complement existing research demonstrating the value of rubrics for feedback and for understanding subject content and assignments. After reviewing some key work on the use of rubrics, the chapter outlines theory around three areas; the importance of context to rubric terminology; the nature of dialogue; the specificity of language to individual assignments to a degree that resists any universal or future transfer

to other assignments. The chapter then provides three examples of potential questions to focus dialogue using rubrics to help students understand what is expected of them in their assignments. The approach is then discussed, specifically in terms of the conditions that can facilitate it, and suggestions made for how others could use it.

Chapter 14

Milena Marinkova, University of Leeds, UK
Joy Robbins, University of Leeds, UK

The growing use of rubrics as tools that can enhance students' learning has prompted an accompanying growth of rubric research in higher education, with a wealth of positive findings. As of yet however, these investigations have predominantly focused on paper-based rubrics or their digital static equivalent rather than truly online rubrics, which present a paradigm shift in how rubrics are displayed, accessed, understood, and interlinked with student text and feedback through the digital affordances of hyperlinking. Studies that have investigated online rubrics so far have focused on pragmatic concerns like efficiency or satisfaction with use, which are important aspects of any digital tool, but secondary to learning. The authors therefore carried out longitudinal case studies to investigate what impacts, if any, the online-ness of rubrics had on students' metacognitive development. Results show strong potential for online rubrics to enhance metacognition, but unfortunately in the majority-used platform we investigated, online rubrics currently are more hindrance than help.

Chapter 15

Martin Sands, King's College London, UK

This chapter highlights key academic contentions around assessing reflection in higher education. Through assimilating experiential, conceptual, and empirical evidence, key terms are defined: 'reflection' as a conceptual and neuroscientific phenomenon within higher education; 'reflective practices' as activities, models, and frameworks that facilitate it; and 'reflective assessments' as content or process driven products of those practices. A literature review elucidates three voices that speak to reflection in higher education.

Chapter 16

Murat Tekin, Çanakkale Onsekiz Mart University, Turkey

High-level cognitive skills are often demonstrated at the performance level. For this reason, performance assessment has become an important element of educational assessment. One of the biggest problems in open-ended questions, oral exams, and performance evaluation is to develop a scoring method that will ensure consistency between raters. At this point, the rubric emerges as a functional scoring tool. Rubrics have many known advantages such as defining the elements and qualities of performance that should be exhibited in educational assessment, consistency between raters, and supporting teaching. On the other hand, there are situations that create barriers in use, such as difficulty in preparation, time consuming to prepare, requiring expertise, and defining the qualifications by distributing them in a balanced way. However, it cannot be said that these are the only obstacles in the use of rubrics. At the same time, educator typology in lesson process and trainer profiles, educational beliefs, and educator roles can be

counted among the important barriers in using rubric.

Chapter 17

Heidi L. Andrade, University at Albany, USA

Rubrics have become ubiquitous in compulsory education and common in higher education. As with any educational innovation, it is time to reflect on the current state of rubrics and how to move ahead. This chapter identifies common conceptions of the rubric that are problematic and proposes redefining the word rubric in terms of learning goals to better align with classroom assessment uses. Feasible suggestions for ensuring the quality of rubrics and avoiding unintended negative consequences for students are also discussed.

Preface

Welcome to *Improving Learning Through Assessment Rubrics: Student Awareness of What and How They Learn,* a comprehensive reference book dedicated to exploring the use and potential of rubrics in the assessment of university learning. This book, edited by Chahna Gonsalves and Jayne Pearson, brings together a diverse range of perspectives, concepts, and practical experiences to shed light on the dynamic and evolving field of learning-oriented assessment.

Educational assessment has long been a focal point in teaching and learning research, with the aim of systematically gathering qualitative and quantitative evidence to evaluate students' progress and achievements. The criteria of dependence, usefulness, reliability, validity, and accuracy guide this process, allowing educators to make informed decisions and provide valuable insights into students' learning journeys. It is within this context that rubrics, as valuable instruments, have emerged to facilitate successful student learning from a competency-based approach.

This book sets out to achieve several key objectives. Firstly, it seeks to identify the strengths and weaknesses of using rubrics for assessing university learning, aiming to improve the overall assessment process. In an ever-changing landscape of higher education, the focus on learning-oriented assessment has become paramount, emphasizing the importance of student progress and academic achievement through formative tasks and continuous assessment.

Throughout the chapters of this book, the editors and contributors delve into various critical aspects related to the use of rubrics in assessment. They explore the psycho-pedagogical implications of rubrics, address reliability and validity concerns in rubric design, discuss the role of students in the assessment process, examine the use of technology for assessment through rubrics, and consider the significance of feedback, particularly "formative feedback," in enhancing students' awareness of their learning process.

Impacting Learning Through Assessment Rubrics is structured to provide theoretical insights, methodological strategies, and practical resources that cater to a wide range of audiences, including educators, researchers, postgraduate students, and education professionals involved in curricular and assessment design responsibilities. The book's scope extends from elementary education to higher education, encompassing vocational education as well.

Moreover, this book strives to contribute to the scientific debate surrounding the use of rubrics in promoting meaningful student learning. By collating various perspectives and evidence-informed practices, it aims to provide an updated and comprehensive overview of the topic, fostering dialogue and encouraging further research in this area.

Assessment is no longer merely a means of assigning grades; it has evolved into a powerful tool that can positively impact learning outcomes. As the field of assessment continues to evolve, this book seeks

to illuminate its participatory and shared nature, emphasizing that assessment is a collective effort aimed at facilitating student learning.

The editors, along with the contributors, recognize that there are ongoing debates, unresolved issues, and potential developments in the use of rubrics for assessing learning. This work, while modest, aims to contribute honest efforts to further scientific knowledge and evidence-based practices concerning rubrics in education.

In conclusion, *Improving Learning Through Assessment Rubrics: Student Awareness of What and How They Learn* serves as a valuable resource for anyone interested in understanding, implementing, and refining the use of rubrics for educational assessment. We hope that this book inspires educators, researchers, and practitioners to engage in continuous improvement and foster a deeper appreciation for the role of rubrics in enhancing the learning experience of students.

CHAPTER OVERVIEW

Section 1: Using Rubrics in Context

This section provides a comprehensive overview of the application of rubrics in various assessment settings, delving into the diverse types of rubrics that are employed across a broad range of disciplines, within higher education institutions, schools, and spanning across different countries. The ensuing chapters equip readers with practical examples and case studies, which shed light on the implementation of rubrics within an assortment of courses and programs. The pivotal role of assessment rubrics in global education becomes apparent as the text explores their significance in assurance of learning and accreditation in higher education. These rubrics offer a lucid, consistent methodology for quantifying and showcasing student achievement and learning outcomes. Assurance of learning pertains to the process that confirms the acquisition of the anticipated knowledge, skills, and competencies by students in a given course or program. On the other hand, accreditation involves the scrutiny of an institution or a program's quality and standards by an external body. Both these processes necessitate evidence of student learning and improvement, which can be systematically and convincingly provided through well-structured and effectively implemented rubrics.

Chapter 1: Using Rubrics for Language Assessment

In this opening chapter, Laura E. Mendoza from the University of Texas, United States, delves into the rapid evolution of rubric usage in classroom assessment, particularly within the context of language classrooms. The chapter highlights the challenge of non-standardized rubrics, which may present language barriers for students and hinder their linguistic development. By offering diverse perspectives on rubrics, the author explores the potential benefits of utilizing rubrics effectively, especially for emergent bilinguals.

Drawing from the insights of Shabani and Panahi (2020), the chapter emphasizes the crucial advantages that arise when language assessment tools, such as rubrics, are introduced to students. These benefits include a better understanding of teacher and rater expectations, improved self-assessment capabilities, enhanced revision skills, and an overall improvement in the quality of students' writing tasks. By

showcasing these advantages, the chapter aims to underscore the significance of incorporating rubrics into language classrooms.

Ultimately, this chapter serves as a call to action for educators, evaluators, and policy-makers to be more mindful when creating and using rubrics in language assessment. By recognizing the potential impact of rubrics on student learning and development, the chapter seeks to empower stakeholders in education to make informed decisions that foster a more inclusive and effective learning environment for all language learners.

Chapter 2: The Use of Rubrics for Drawing Graphs in Physics Education

Authored by Işıl Aykutlu from Hacettepe University, Turkey, Chapter 2 focuses on the application of rubrics in the context of physics education, specifically concerning the drawing of graphs. In physics classes, graphs play a crucial role in enhancing students' understanding of the relationships between variables. They serve as valuable tools, particularly in demonstrating the connections between dependent and independent variables in laboratory experiments.

Drawing an accurate and meaningful graph requires students to possess a solid understanding of physics concepts, as well as the ability to interpret and depict the characteristics of graphs. To achieve this, they must follow various sub-steps in the graph-drawing process. This chapter emphasizes the importance of assessing students' graph drawings using rubrics, as it offers a more detailed and reliable evaluation of their performance, making the assessment process more valid.

Furthermore, the chapter highlights the benefits of incorporating rubrics into the teaching process itself. By using rubrics to instruct students on how to draw graphs, educators can contribute to their comprehension of graph characteristics and aid in clearly outlining the necessary steps. Moreover, rubric-based instruction encourages self-assessment and peer assessment, providing students with opportunities for reflection and improvement in their graph-drawing abilities.

Through the utilization of rubrics, this chapter suggests that educators can foster a more effective and comprehensive learning experience in physics education. The application of rubrics in assessing and instructing graph drawing not only enhances the reliability of evaluations but also contributes to students' deeper understanding of the subject matter, ultimately empowering them to become proficient graph-drawers and critical thinkers in the realm of physics.

Chapter 3: Fostering Entrepreneurship Education by Improving Assessment Rubrics for Entrepreneurship Competence

Authored by Minna Hämäläinen and Anu Raappana from Lappeenranta-Lahti University of Technology, Finland, Chapter 3 centers on a comprehensive tool designed to assess learners' entrepreneurship competence and its creation process. Entrepreneurship competence is a crucial skill promoted at all levels of education, but the lack of adequate assessment tools poses a challenge for teachers. This chapter addresses this issue by introducing a rubric-based tool that facilitates the evaluation of students' entrepreneurship competence in a straightforward and accessible manner.

The rubric-based assessment tool not only empowers teachers to gauge the development of learners' entrepreneurship competence but also provides clarity on what aspects should be assessed and when to conduct assessments. By utilizing the rubric, teachers can gain insights into the specific dimensions of

entrepreneurship competence that are being evaluated, thus further enhancing their own professional skills in this domain.

One notable aspect of the tool presented in this chapter is its inclusivity and international scope. The development process involved a participatory approach, with collaboration among teachers, students, and researchers. This inclusive method ensures that the rubric is adaptable and applicable in diverse cultural contexts, promoting a holistic understanding of entrepreneurship competence and its assessment across different settings.

As educators strive to foster entrepreneurship education and cultivate essential skills in their students, the rubric-based assessment tool outlined in this chapter serves as a valuable resource. By providing a structured and effective means of assessing entrepreneurship competence, it not only empowers teachers but also facilitates students' growth and development in this critical area. Additionally, the international perspective of the tool encourages cross-cultural contributions, enriching the overall learning experience and fostering a deeper appreciation for entrepreneurship education on a global scale.

Chapter 4: Learning how to Become a Teacher Researcher: Using Rubrics to Support Evidence-Informed, Research-Based Practice

Authored by Emma O. Brien, Josephine Brady, T. J. Ó Ceallaigh, Katherine Babbitt, Andrea Brosnan, Emma Byrne, Erin Byrne, Rebecca Curtin, Lisa Gaffney, and Karen O. Callaghan, all affiliated with Mary Immaculate College, Ireland, Chapter 4 delves into the vital aspect of developing teachers as reflective practitioners within Irish teacher education. While the importance of reflective practice is acknowledged, a significant gap often exists between theoretical principles and practical implementation.

This chapter sheds light on the positive impact of rubrics on student self-reflection and self-regulation, offering valuable insights into their application within the framework of supervisory relationships. Drawing upon Drytons' extended Supervisory Working Alliance, the authors explore how rubrics can foster productive student-supervisor relationships during the research process.

Employing a Collaborative AutoEthnographic (CAE) methodology, the researchers authentically capture both student and faculty perspectives, providing a comprehensive and well-rounded understanding of the subject matter. The findings demonstrate that rubrics play a pivotal role in promoting transparency and identifying clear expectations for students. They provide a common language for students to articulate and evaluate their work, enabling them to take ownership of the feedback process and actively engage in discussions with their supervisors.

Moreover, rubrics serve as a metadialogue, facilitating bonding between students and supervisors while shaping meaningful conversations about the research work. By offering a contextual framework for feedback, rubrics empower students to comprehend and apply constructive feedback effectively, ultimately fostering evidence-informed and research-based practice.

This chapter offers valuable insights for educators and teacher training institutions seeking to bridge the gap between theoretical concepts and practical implementation of reflective practice. The application of rubrics as a supportive tool in teacher research can significantly enhance the supervisory relationship, promote reflective learning, and contribute to the overall growth and development of teacher researchers. As we explore this innovative approach, we hope to inspire further research and exploration of rubrics as a facilitator of evidence-informed, research-based practice in teacher education.

Chapter 5: The Use of a Rating Scale as a Formative and Shared Assessment Tool in Physical Education

Authored by Daniel Bores-García, Raúl A. Barba-Martín, Gustavo González-Calvo, and David Hortigüela-Alcalá, affiliated with various Spanish universities, Chapter 5 explores a formative and shared assessment experience in the context of Physical Education at the Secondary Education level in a high school in Spain.

The chapter focuses on the utilization of an assessment scale as an assessment instrument in the evaluation process. Through this rating scale, students actively participate in the assessment process by engaging in self-assessment and co-assessment with their peers. Additionally, the teacher provides an external evaluation based on criteria established in the instrument.

After implementing this formative and shared assessment approach, a study was conducted to explore students' perceptions of the process and their experience with the evaluation scale. The researchers employed various methods, including discussion groups and the analysis of teachers' and students' diaries, to gain deeper insights into the impact of this assessment method.

The results of the study reveal that students exhibited heightened motivation and commitment to both the task at hand and the overall group dynamic. This positive experience was attributed to the sense of ownership and involvement that students felt by being integral participants in the assessment process. The collaborative nature of the formative and shared assessment method fostered a sense of belonging and investment in the learning journey.

By showcasing the benefits of this assessment approach, the chapter advocates for the wider adoption of formative and shared assessment practices in Physical Education and other educational settings. The incorporation of rating scales as assessment tools empowers students to take charge of their learning, develop self-awareness, and engage in constructive evaluations, leading to a more enriched and meaningful learning experience.

Overall, this chapter sheds light on the transformative impact of formative and shared assessment in Physical Education and offers valuable insights for educators seeking to create more inclusive, collaborative, and student-centered learning environments. Through this exploration, we hope to inspire educators to embrace innovative assessment practices that empower students as active participants in their educational journey, ultimately fostering a positive and rewarding learning experience.

Chapter 6: Effective Use of Rubrics in Student Evaluation: Best Practice e-Portfolios

Authored by Elena Ramona Richiteanu-Nastase and Alexandru Robert Mihaila from Bucharest University of Economic Studies, Romania, Chapter 6 offers a comprehensive exploration of the effective utilization of rubrics in student evaluation, specifically focusing on best practice e-portfolios as an alternative evaluation method.

The chapter commences by elucidating essential concepts related to evaluation, including traditional and alternative evaluation methods, process-centered evaluation, and the evaluation of student progress. Portfolios are introduced as a noteworthy alternative evaluation method. The authors then delve into the concept of rubrics as a valuable and versatile evaluation tool, providing insights into their design and exemplification for commonly used evaluation methods. The advantages and disadvantages of rubrics are thoroughly analyzed, considering the most important issues and controversies surrounding their implementation.

In the subsequent sections, the chapter delves into a practical example of good practice—the use of rubrics in student evaluation through e-portfolios. E-portfolios are explored as an innovative and comprehensive method of assessment that complements rubric-based evaluation. The authors shed light on how this combination can enhance the evaluation process, offering a more holistic view of students' progress and achievements.

The final part of the chapter is dedicated to discussions and recommendations for using rubrics in evaluation. The authors critically examine the utility of rubrics and their limitations as an assessment instrument. They also explore avenues for further development and improvement of rubric-based evaluation practices.

Overall, this chapter serves as a valuable guide for educators and institutions seeking effective ways to assess students' progress and performance. By emphasizing the advantages and potential drawbacks of rubrics and e-portfolios, the authors provide a comprehensive understanding of their combined impact in student evaluation. Through thoughtful analysis and practical examples, this chapter inspires educators to adopt best practices in using rubrics and e-portfolios for student assessment, ultimately contributing to a more robust and student-centered evaluation process.

Section 2: Developing Rubrics

This section navigates through the dynamic process of developing rubrics for learning and assessment with colleagues and students. Collaborative development provides a rich, engaging environment that draws on the collective wisdom and experiences of a diverse team. To do this, the section presents a comprehensive guide for engaging in productive collaboration, articulating steps to establish a clear purpose, foster open dialogue, and encourage the active participation of all involved parties. It further elucidates diverse modes of collaborative processes and their respective evaluations, catered to different contexts, reflecting the diversity and variability of educational environments. This is important as it democratizes the learning process, promotes ownership, and enriches the content with multiple perspectives. Moreover, it fosters a sense of collective responsibility, enhancing commitment towards achieving educational goals. The value of this process also lies in its potential to generate deeper understanding and empathy among educators and learners, thereby creating a more inclusive and responsive learning environment. Thus, this section is not just a practical guide for collaboration, but also a call to action for educators to embrace the transformative potential of joint development in enhancing the quality and relevance of education.

Chapter 7: The Importance of Student Partnership in Rubric Construction: Discussion and Evaluation

Authored by Allan Stephen Laville, Lindsey Thompson, Yue Yue, Alexandra J. Hayward, and Victoria Grace-Bland from the University of Reading, United Kingdom, Chapter 7 delves into the significance of student-staff partnerships in the development and evaluation of rubrics. Anchored in the University of Reading's principles for student-staff partnerships, this chapter places student voice at the core of the process to enhance teaching and learning initiatives.

The chapter delves into the challenges of engaging students with assessment rubrics and highlights the efforts taken to overcome these barriers through active listening exercises. By incorporating stu-

dent perspectives, the authors aim to foster greater student engagement and participation in the rubric development process.

To illustrate the practical implications of student-staff partnerships in rubric construction, the chapter presents three case studies. These case studies offer valuable insights and recommendations for improving student assessment literacy. They advocate for in-class support for rubrics, including interactive discussions to clarify rubric criteria and expectations. Additionally, the chapter emphasizes the importance of supplementary support mechanisms outside the classroom, such as assessment rubric screencasts and discussion boards, to provide further guidance and enhance students' understanding of rubrics.

One essential aspect highlighted in the chapter is the concept of co-creation in developing new rubrics. By involving students in the construction of rubrics, educators can better align assessment criteria with students' needs and expectations, leading to more effective and meaningful evaluation processes.

In conclusion, the chapter emphasizes the critical role of student-staff partnerships in the development and evaluation of rubrics. By actively involving students in the process, educators can enhance student engagement, understanding, and ownership of the assessment process. The chapter also underscores the importance of establishing additional support mechanisms to nurture both student and staff assessment literacy effectively.

Overall, this chapter serves as a valuable resource for institutions and educators seeking to create more inclusive and student-centered assessment practices. By adopting student-staff partnerships and co-creation in rubric development, educators can foster a positive learning environment that empowers students and promotes their active participation in the assessment process.

Chapter 8: Co-Production of Assessment Rubrics in an Online Education Context

Authored by Anja Harrison, Maren Beier, Harriet Power, and Brenda P. Williams from King's College London, United Kingdom, Chapter 8 delves into the significance of co-production in developing assessment rubrics in the context of online education. Marking rubrics are highly regarded for their transparency and efficacy in supporting student success, facilitating their understanding of the learning environment and helping them achieve their academic goals. Rubrics also aid staff in fair and consistent grading. However, for rubrics to be successful, they must be accessible to all stakeholders, necessitating active engagement from both students and staff.

The chapter emphasizes the importance of creating rubrics that are comprehensible and inclusive to a diverse student body, particularly in the context of online courses where cultural diversity is often more pronounced. Understandability is key, and inclusive wording is crucial in accommodating the varying backgrounds and experiences of students. Co-creation of rubrics with students becomes an effective approach to fostering inclusivity and developing meaningful and successful rubrics.

To provide practical guidance on co-creating and implementing rubrics in an online education context, the chapter presents a step-by-step guide. This guide has been developed through collaboration with the co-creation student panel from the online programs at the Institute of Psychiatry, Psychology, and Neuroscience, KCL.

By involving students in the rubric development process, educators can better align the assessment criteria with students' needs and perspectives, leading to more equitable and effective evaluation practices. The chapter showcases how co-production of rubrics not only enhances the understandability and relevance of assessment criteria but also fosters a sense of ownership and empowerment among students.

In conclusion, the chapter advocates for the integration of co-production practices in the creation of assessment rubrics in online education settings. By leveraging the expertise and insights of both students and staff, institutions can develop rubrics that promote inclusivity, understanding, and meaningful evaluation for a diverse student population.

Overall, this chapter serves as a valuable resource for educators and institutions seeking to enhance assessment practices in online education. By embracing co-creation principles and incorporating student input, educators can create more accessible and student-centered rubrics that empower learners to excel in their academic journey.

Chapter 9: Facilitating Program Level Assessment Working Teams to Develop Shared Rubrics Across an UG and PG Program in Business Education

Authored by Roisin Donnelly and Colin Hughes from Technological University of Dublin, Ireland, Chapter 9 presents a reflective study on a college-wide initiative undertaken in the Faculty of Business at a Technological University in Ireland. Focusing on the critical area of assessment and feedback, the chapter highlights the collaborative nature of the initiative, its broad scope, and the extensive support provided to both staff and students throughout the programmatic assessment design process.

The chapter centers on the development of rubrics as the primary focus, with an overview of Program Learning Outcome (PLO) mapping provided as context. To facilitate this process, four rubric working groups were formed across the College, each assigned to develop common rubrics for specific areas: reflective practice, critical thinking, presentation skills, and industry consultancy projects. The work of these rubric groups was an integral part of the college-wide PLO mapping project.

This program-based study delves into the challenges of bridging cultures, practices, and understandings within disciplinary teams in a technological university context. It emphasizes the complexities of bringing together different perspectives and experiences to create shared rubrics that align with the program's learning outcomes and goals.

The chapter showcases the significance of collaboration and mutual understanding among faculty members, staff, and students in the assessment design process. It underscores the importance of support mechanisms in fostering a conducive environment for the successful development of shared rubrics across undergraduate and postgraduate programs in business education.

Ultimately, this chapter provides valuable insights into the practical implementation of programmatic assessment and rubric development within a technological university setting. It highlights the collective effort and commitment required to establish cohesive and effective assessment practices that align with the institution's overarching goals and promote student success.

In conclusion, the chapter encourages educators and institutions to embrace a collaborative and supportive approach to assessment and feedback design. By fostering interdisciplinary partnerships and engaging in ongoing dialogue, institutions can create shared rubrics that enhance the learning experience for students and contribute to the continuous improvement of business education programs.

Chapter 10: Developing a Rubric to Evaluate Dissertations
in Educational and Social Sciences

Authored by Ömer Açıkgöz, Aydin Aslan, Korkut Koçak, Aslı Günay, and Nevzat Yavuz from various universities and institutions in Turkey, Chapter 10 presents a study focused on developing a rubric for the evaluation of dissertations conducted in the fields of educational and social sciences.

The chapter outlines the key objectives of the study, aiming to create a comprehensive rubric that considers the knowledge, skills, and competencies required at the doctoral level within the European and Turkish Qualifications Frameworks. Additionally, the rubric development process takes into account the legal framework of Turkish Higher Education and incorporates the perspectives of 12 experts in the fields of educational and social sciences. By integrating these diverse elements, the authors aim to create a robust and reliable rubric that can effectively evaluate dissertations within these academic disciplines.

The rubric's significance lies in its potential to contribute to the evaluation of completed dissertations in the educational and social sciences fields, providing a structured approach to assess academic studies based on empirical evidence. As such, the rubric serves as a valuable guide for PhD students, researchers, and academics, facilitating a more systematic and standardized evaluation process.

By utilizing this rubric, institutions and academic communities can establish consistent and objective evaluation criteria, fostering quality assurance and promoting the advancement of research within educational and social sciences disciplines.

In conclusion, Chapter 10 highlights the importance of developing a well-structured and inclusive rubric for assessing dissertations in educational and social sciences. The collaborative approach, involving experts and considering various frameworks and legal aspects, ensures the rubric's comprehensive nature and its potential to elevate the quality of research and academic output in these fields. This contribution to the field of educational and social sciences holds promise for advancing doctoral studies and furthering research excellence in Turkey and beyond.

Section 3: New Perspectives on Rubrics

This portion of the book pushes the boundaries of understanding rubrics by introducing fresh perspectives and new contexts, fostering a discussion about their limitations, potential, and future directions. The third section delves into the profound impact of rubrics on student learning, motivation, and faculty professional development, with particular emphasis on their role in evaluating reflective practice.

In the field of teacher development, the section examines the significance of rubrics in fostering both teacher and student assessment literacy. Assessment rubrics, when effectively designed and applied, offer potent solutions to enhance assessment and feedback practices, an area often highlighted for improvement in student experience surveys like the UK's National Student Survey (NSS) and the Postgraduate Taught Experience Survey (PTES). The 2022 NSS results confirmed the urgent need for reform in this area, with only 69% of students expressing satisfaction.

The value of rubrics extends beyond immediate assessment practices, with the potential to boost student metacognition. This gives students more control over their own learning process, leading to improved outcomes. The pedagogical potential of rubrics is significant, despite some didactic limitations, and their use as feedforward tools provides opportunities for meaningful, subject contextualized dialogue.

While acknowledging the profound implications for teacher development and training, the section also anticipates the future of rubrics, encouraging further exploration and experimentation to unlock

their full potential in a rapidly evolving educational landscape. This exploration will pave the way for more strategic and impactful use of rubrics, ultimately enhancing both student learning and satisfaction.

Chapter 11: Knowledge of Language in Rubric Design: A Systematic Functional Linguistics Perspective

Authored by Chahna Gonsalves from King's College London, United Kingdom, Chapter 11 delves into the critical but often overlooked aspect of language choice in rubric design. While rubrics have gained popularity as valuable tools for assessment and instruction in higher education, their efficacy, particularly in instructional and formative contexts, can significantly depend on the language used within them.

The chapter commences by providing a comprehensive review of current research and guidance on effective rubric language. It highlights the importance of considering the implications of linguistic choices on students' awareness of their learning process and outcomes. Developing an effective rubric that guides students' understanding of what and how they learn requires thoughtful consideration of language constructs.

To enhance teachers' capacities in effective rubric design, the author introduces the theory of systemic functional linguistics (SFL). This linguistic framework offers valuable insights into how language functions in different contexts and can be employed to foreground language considerations in rubric design. By integrating SFL-informed training, teachers can improve their rubric development skills, resulting in more effective and student-friendly rubrics.

The chapter emphasizes the vital role of language knowledge in both teacher and student literacy development. By enhancing teachers' understanding of language constructs and their implications, educators can better foster academic and assessment literacy in students. This, in turn, enables students to comprehend assessment criteria, engage with feedback, and take ownership of their learning journey.

Ultimately, the chapter underscores the significance of developing teachers' knowledge about language and its impact on rubric design. Empowering educators with language-centered training can lead to the creation of more meaningful and effective rubrics, enhancing the assessment and instructional experience for both teachers and students.

In conclusion, Chapter 11 sheds light on the interplay between language and rubric design, providing valuable insights for educators seeking to improve their assessment practices. By embracing the systematic functional linguistics perspective and incorporating language considerations in rubric design, educators can promote student learning and literacy development, contributing to a more enriching and impactful educational experience.

Chapter 12: Self-Assessment: Preservice Teachers' Concepts, Instruments, and Practices

Authored by Elsa Maria Ferro Ribeiro-Silva and Catarina Amorim from the University of Coimbra, Portugal, Chapter 12 delves into the critical aspect of self-assessment in students' learning. The chapter presents the findings of a study conducted with 72 university students, focusing on preservice teachers' perspectives on self-assessment, the instruments they use, and the implications for their students' learning.

The chapter highlights the significance of self-assessment as a pivotal form of assessment for promoting student learning. While preservice teachers demonstrate a theoretical understanding of self-assessment, the study reveals that in practice, it is often conducted in a non-systematic manner with quantitative

criteria, typically planned on a designated day. This suggests a disconnect between the theoretical understanding and actual implementation of self-assessment.

The results indicate that preservice teachers may view self-assessment as a legal requirement rather than an essential tool for students to critically reflect on their learning. The chapter sheds light on the challenges and implications of this misalignment between understanding and practice, potentially hindering the full potential of self-assessment in supporting students' growth and development.

By examining preservice teachers' perspectives and practices related to self-assessment, the chapter offers valuable insights into the need for further training and support in this area. Emphasizing the importance of fostering a deeper understanding of self-assessment, educators can better equip preservice teachers with the necessary tools and strategies to effectively implement self-assessment as a valuable and meaningful assessment practice.

In conclusion, Chapter 12 calls attention to the gap between theoretical knowledge and practical application of self-assessment in preservice teachers' practices. The findings underscore the need for a more thoughtful and intentional approach to self-assessment, one that goes beyond mere compliance with legal requirements. By promoting critical reflection and student-centered assessment practices, educators can harness the full potential of self-assessment as a powerful tool for enhancing students' learning experiences and fostering their growth and development.

Chapter 13: Using Rubrics as Feedforward Tools for Subject Contextualized Dialogue

Authored by Kendall Richards and Nick Pilcher from Edinburgh Napier University, United Kingdom, Chapter 13 presents an innovative approach to using rubrics as feedforward tools in a subject contextualized dialogue between teachers and students. This process occurs before students submit their assignments, aiming to complement existing research on the value of rubrics for feedback and enhancing students' understanding of subject content and assignments.

The chapter begins by reviewing key research on the use of rubrics, establishing their importance in the assessment and feedback process. It then delves into three crucial areas of theory: the significance of context in rubric terminology, the nature of dialogue between teachers and students, and the specificity of language tailored to individual assignments, making it challenging to universally transfer rubric usage to future assignments.

Subsequently, the chapter provides three examples of potential questions that can guide the dialogue using rubrics to help students grasp what is expected of them in their assignments. This approach highlights the role of rubrics as tools for feedforward, facilitating meaningful discussions between teachers and students to enhance the students' understanding of assignment requirements and expectations.

The chapter further examines the conditions that facilitate this feedforward process and discusses how educators can apply this approach effectively. By engaging in contextualized dialogue using rubrics, teachers can provide tailored guidance and support to students, leading to improved assignment outcomes and learning experiences.

In conclusion, Chapter 13 emphasizes the value of using rubrics as feedforward tools for subject contextualized dialogue. By engaging in proactive discussions with students before the assignment submission, educators can enhance students' comprehension of assignment expectations and encourage deeper engagement with subject content. The chapter's practical examples and suggestions offer educators valuable insights into implementing this approach in their teaching practices, ultimately benefiting both teachers and students in the assessment and learning process.

Chapter 14: Can Online Rubrics Develop Learners' Metacognition: A Qualitative Case Analysis Study

Authored by Milena Marinkova and Joy Robbins from the University of Leeds, United Kingdom, Chapter 14 explores the potential of online rubrics to enhance students' metacognitive development. As the use of rubrics as learning tools continues to grow in higher education, research has mostly focused on paper-based or static digital rubrics. However, online rubrics present a paradigm shift in how rubrics are displayed, accessed, and understood, utilizing digital affordances like hyperlinking to interlink with student text and feedback.

The chapter highlights the importance of investigating the impact of online rubrics on students' metacognitive development, which refers to their ability to reflect on and regulate their learning processes. While previous studies have mainly focused on pragmatic aspects like efficiency and user satisfaction, this study delves deeper to explore the potential of online rubrics to enhance metacognitive skills.

To investigate this, the authors conducted longitudinal case studies, aiming to understand how online rubrics influenced students' metacognition. The results revealed promising potential for online rubrics to enhance metacognitive development. However, the study also uncovered challenges in the current platforms used, where online rubrics were found to be more of a hindrance than a help.

The chapter sheds light on the need for further exploration and improvement in the design and implementation of online rubrics to fully harness their potential as metacognitive development tools. By leveraging the digital affordances and interactivity of online rubrics, educators can facilitate students' reflective thinking, self-assessment, and self-regulation, ultimately fostering more effective and independent learners.

In conclusion, Chapter 14 calls attention to the untapped potential of online rubrics in developing students' metacognitive skills. The findings underscore the significance of intentional design and research to ensure online rubrics truly support and enhance metacognitive development. By addressing the current challenges and focusing on the transformative impact of digital affordances, educators can pave the way for more effective use of online rubrics to foster students' metacognitive growth, ultimately empowering them as self-directed learners.

Chapter 15: Exploring and Developing Reflective Writing Rubrics in Higher Education

Authored by Martin Sands from Kings College London, United Kingdom, Chapter 15 delves into the complexities of assessing reflection in higher education. The chapter focuses on key academic debates and contentions surrounding the evaluation of reflective writing in the context of higher education.

To build a comprehensive understanding, the author assimilates experiential, conceptual, and empirical evidence to define essential terms related to reflection. "Reflection" is explored as a conceptual and neuroscientific phenomenon within higher education, highlighting its significance in the learning process. "Reflective practices" are identified as the activities, models, and frameworks that facilitate the process of reflection, providing students with opportunities to engage in critical self-analysis and learning from their experiences. "Reflective assessments" are then explained as the content or process-driven products of those reflective practices, emphasizing their role in evaluating students' reflective abilities and growth.

The chapter proceeds with a thorough literature review, uncovering three distinct perspectives or voices that contribute to the understanding of reflection in higher education. These perspectives offer

valuable insights into different approaches to assessing reflection and highlight the diverse academic viewpoints on the topic.

By addressing key academic contentions and synthesizing evidence from various sources, the chapter lays the groundwork for a deeper understanding of reflective writing rubric development. Through this exploration, educators and researchers gain a more comprehensive understanding of how to effectively assess and foster students' reflective abilities in higher education.

In conclusion, Chapter 15 provides a valuable contribution to the field of higher education assessment by shedding light on the complexities of assessing reflection. The synthesis of experiential, conceptual, and empirical evidence, along with a review of multiple perspectives, offers a well-rounded understanding of reflective writing rubric development. This understanding can inform educators' practices in fostering students' critical thinking, self-awareness, and metacognitive skills, ultimately enriching their learning experiences in higher education.

Chapter 16: Pedagogical Potential and Didactic Limitations of Assessment Rubrics: An Example from Medical Education

Authored by Murat Tekin from Çanakkale Onsekiz Mart University, Turkey, Chapter 16 explores the pedagogical potential and didactic limitations of assessment rubrics, focusing on their application in the context of medical education. High-level cognitive skills are often demonstrated through performance-based assessments, making performance assessment a critical component of educational evaluation.

The chapter highlights a major challenge in open-ended questions, oral exams, and performance evaluations: the need for a scoring method that ensures consistency between raters. In this context, rubrics emerge as functional scoring tools that offer several advantages, including defining the elements and qualities of performance expected in educational assessment, ensuring consistency between raters, and supporting teaching and learning processes.

Despite their benefits, the use of rubrics also presents certain barriers. One notable challenge is the difficulty in preparing rubrics, which can be time-consuming and require expertise to define the necessary qualifications and distribute them in a balanced manner. However, the chapter emphasizes that these are not the only obstacles to rubric usage.

The educator typology, including teaching style and trainer profiles, as well as educational beliefs and educator roles, can also play significant roles in hindering the effective use of rubrics. Understanding and addressing these factors are crucial in overcoming barriers and maximizing the potential of rubrics in medical education.

In conclusion, Chapter 16 sheds light on the importance of assessing the pedagogical potential and didactic limitations of rubrics in the context of medical education. By acknowledging the challenges and understanding the unique dynamics within the educational environment, educators can implement rubrics more effectively to support performance-based assessment and enhance the learning experience for medical students. This chapter serves as a valuable resource for educators seeking to optimize the use of rubrics and foster a more consistent and impactful educational assessment process in medical education.

Chapter 17: What is Next for Rubrics: A Reflection on Where We Are and Where to Go from Here

Authored by Heidi L. Andrade from the University at Albany, United States, Chapter 17 offers a reflective examination of the current state of rubrics and explores future directions for their use in education. Rubrics have become widely adopted in compulsory education and are common tools in higher education. However, as with any educational innovation, it is essential to assess the effectiveness and impact of rubrics to ensure they align with evolving pedagogical needs.

The chapter begins by identifying common conceptions of rubrics that may be problematic and hinder their full potential as assessment tools. By redefining the concept of rubrics in terms of learning goals, the author proposes a more meaningful alignment with classroom assessment uses. This reframing encourages educators to view rubrics as dynamic tools that support and guide students' progress toward specific learning objectives.

Furthermore, the chapter delves into ensuring the quality of rubrics and preventing unintended negative consequences for students. As rubrics are widely used in educational settings, it is crucial to establish best practices for creating and implementing them effectively. Attention is given to addressing potential issues related to fairness, clarity, and consistency in rubric design and application.

In conclusion, Chapter 17 provides educators with an insightful reflection on the current state of rubrics and the direction they should take in the future. By redefining rubrics to align with learning goals and establishing quality assurance measures, educators can leverage these tools to foster more meaningful and effective assessment practices. The chapter encourages educators to embrace continuous improvement in rubric development and usage, ultimately enhancing their potential to support student learning and growth in educational settings.

IN SUMMARY

As we conclude this edited reference book, "Improving Learning Through Assessment Rubrics: Student Awareness of What and How They Learn," we reflect on the journey that brought us together to explore the vast landscape of assessment in education. Throughout this book, we have delved into the profound impact of rubrics as powerful tools for guiding student learning and enhancing the assessment process.

Our aim in compiling this volume was to provide a platform for a critical dialogue surrounding the use of rubrics in evaluating university learning. We sought to uncover the theoretical underpinnings, methodological strategies, and practical experiences that contribute to the meaningful implementation of rubrics in educational settings. As editors, we are humbled by the wealth of knowledge shared by our esteemed contributors, each offering unique insights into this evolving field.

The book has taken us on a journey through the diverse perspectives and contexts in which rubrics are utilized, ranging from elementary to higher education and vocational settings. We have explored the relevance of rubrics in the current educational landscape, emphasizing that assessment should no longer be confined to mere grading but rather be an integral part of the learning process itself.

In a world of rapid changes and advancements in education, the concept of learning-oriented assessment has emerged as a pivotal approach to promote student progress and success. Rubrics, as guiding instruments, offer a roadmap for both educators and students, enhancing clarity on learning expectations and desired outcomes. By providing constructive feedback and facilitating feedforward, rubrics have

proven to raise students' awareness of their learning journey and foster a deeper understanding of how to apply their knowledge to future situations.

Through this book, we have endeavored to contribute to the ongoing scientific debate on the quality and effectiveness of rubrics in educational assessment. By shedding light on both the potential and limitations of rubrics, we hope to inspire further research and evidence-informed practices in this domain.

Our gratitude goes out to all the contributors who generously shared their expertise, experiences, and passion for improving the assessment process. We also extend our appreciation to the teachers, researchers, postgraduate students, and education professionals who will read and engage with this book. It is through your dedication and commitment to quality education that meaningful changes can be brought about in the assessment landscape.

As we bid farewell to this compilation, we remain optimistic that "Improving Learning Through Assessment Rubrics" will serve as a valuable resource for years to come. May the knowledge disseminated in these pages foster innovative approaches to assessment and contribute to the holistic growth and development of students worldwide.

Let this book be a beacon, guiding educators and learners alike towards a brighter future, where assessment is not only a measure of achievement but a catalyst for continuous improvement and lifelong learning.

Chahna Gonsalves
King's Business School, King's College London, UK

Jayne Pearson
King's Academy, King's College London, UK

Acknowledgment

Editorial Advisory Board

Daniel Drumm, *King's College London, UK*
Chahna Gonsalves, *King's College London, UK*
Zhonghan Lin, *King's College London, UK*
Andrew McFaull, *King's College London, UK*
Jayne Pearson, *King's College London, UK*

Section 1
Using Rubrics in Context

An overview of how rubrics are used- the different types employed for different assessment settings. This covers a range of disciplines, HE and schools, and different countries. This section contains much of the introductory descriptive stuff about rubrics and is best placed first.

Chapter 1
Using Rubrics for Language Assessment

Laura E. Mendoza
iD https://orcid.org/0000-0001-8649-8775
University of Texas at El Paso, USA

ABSTRACT

The use of rubrics for classroom assessment has been evolving rapidly during the last decades. Unfortunately, because rubrics across language classrooms are only sometimes standardized, some may provide a non-friendly language for students, which is useless for their linguistic development. In addition to defining rubrics from an array of perspectives, the present chapter presents possible benefits linked to the appropriate use of rubrics among emergent bilinguals. Shabani and Panahi highlight essential benefits for students when presented with language assessment tools, such as rubrics, and the authors use their observation for this chapter. It aims to highlight the importance of rubrics in the language classroom in an effort for educators, evaluators, and policymakers to be more conscious when creating and using rubrics.

INTRODUCTION

Assessing and, more importantly, addressing students' learning can be challenging in any classroom. Particularly in the language classroom, assessing students' needs can represent additional challenges given the many capabilities that, interposingly, are continuously present. It is relevant to remind the reader that although the word 'rubric' is continuously used throughout the chapter, the concept is to be mainly used in the language classroom context. Nevertheless, readers should be warned that the term 'rubric' is used today in arrayed contexts and practices. However, the focus of the chapter will remain as an evaluative tool to assess students' linguistic practices and knowledges. The aim of the present chapter is to highlight the importance of rubrics in the language classroom in an effort for educators, evaluators, and policy-makers to be more conscious when creating and using rubrics.

BACKGROUND: DEFINING RUBRICS AND TYPES OF RUBRICS

The use of rubrics for classroom assessment has been evolving rapidly during the last decades. In education, for instance, teachers always use rubrics even without noticing it. Jeong (2015) reminds us how the use of rubrics, particularly in the language classroom, is of high value. For the purpose of the present chapter, a rubric should be understood as presented by Dawson (2017, p. 349):

A rubric is a tool used to assess student work that usually includes Popham's (1997) three essential features: evaluative criteria, quality definitions for those criteria at particular levels, and a scoring strategy. A design element is a particular variable, choice, or dimension that makes one rubric different from another. For example, the specificity element is concerned with the differences between task-specific and generic rubrics.

According to the author, "one rubric may use generic quality words (e.g., 'good' or 'below standard'), whereas another may explain in detail what quality looks like" (p. 348). This said the level of specificity that a rubric may possess will serve as a guide for students, educators, and evaluators to gain a better understanding of the course expectations.

Some educators unknowingly use mental rubrics (Quinlan, 2012), whereas others purposefully use physical/digital rubrics. The author states: "Whether they know it or not, people create rubrics - guidelines for decisions for evaluation and assessment- in their minds every day" (p.2). Quinlan (2012) continues, "These mental rubrics help us to make decisions based on both our prior knowledge and current objectives." In our classrooms, we generally use rubrics with the presumption that these will ease the path for students when trying to understand the expectations, components, objectives, and ways to be assessed in a given course or assignment; we are continuously bringing opportunities for students, and ourselves, to compare what it is good or not that good.

It has been evident that, particularly for students, rubrics in the language classroom can help them successfully prove what they are expected to do linguistically if provided with details. Dawson (2017) argues that there is a relevant variation amongst the level of detail provided in the quality section of the rubrics; therefore, this opens room to having rubrics that can be helpful for students and evaluators, whereas others need to be more specific. These distinctions will be thoroughly discussed later on as part of the present chapter. Therefore, the use of rubrics in the language classroom can be beneficial to demonstrate student success and for educators and evaluators to provide more detailed-oriented feedback, promoting meaningful learning.

Unfortunately, because rubrics across language classrooms are not always standardized, some of them may provide a non-friendly language for students, which is useless for their linguistic development; likewise, many others provide a language that might be too vague to comply with institutions' missions and objectives. Dawson (2017) highlights how the proliferation of the word 'rubric' has dealt with varied changes and adaptations. However, simultaneously, it has yet to develop a shared complete understanding of the term necessarily. For this reason, although many institutions require using rubrics for varied assessments, these fluctuate, navigating the use of the term and using it in the way that best benefits them.

Commonly Used Types of Rubrics

Rubrics can vary depending on many factors. In the educational world, this can vary, for example, concerning the educational grade level, type of assignment, type of specificity, and type of subject, to name a few. Nevertheless, although rubrics can be presented in varied forms and sizes, two distinguishable elements are always present: They present performance levels, such as producing quality work, and those of quality, usually ranking from below expectations to outstanding (Glass, 2004). Generally speaking, rubrics tend to include three-to-six performance criteria, as these are genuinely manageable for the scorer to remember (Wolf & Stevens, 2007).

Rubrics are, for instance, divided according to their performance or the quality they are presupposed to assess, as previously mentioned. The performance factors can be general or written for a specific assignment. Quality levels can be numerical, in word form, or a combination of both (Glass, 2004, p. 17). Some other rubrics can range according to their levels of precision when assessing. Commonly, many other rubrics are based on the "original six-trait scoring guide developed by teachers for teachers through the efforts of the Six-Trait Analytical Scoring Committee of the Beaverton, Oregon, School District. […] The categories of the six traits are as follows: Ideas, organization, voice, word choice, sentence fluency, conventions" (Glass, 2004, p. 19).

These categories allow educators to quantify and therefore highlight the impact of the traits students are commonly demonstrating when completing a task; with the use of these rubrics, educators are able to quickly illustrate if students' ideas were developed appropriately and if their tone of voice was adjusted to the audience. In the language classroom, these categories can inform educators and students if language was precisely used in terms of word choice. For example, equally important, these rubrics can facilitate the dissemination of standard grammar conventions in a particular language or context. Research has demonstrated that when using the six-traits rubric as an instruction and assessment tool, students show higher performance, primarily in writing (Glass, 2004). In a similar fashion, other authors (Lee et al., 2010) have noted that, generally, these traits' scores tend to correlate among themselves.

Rubrics can also be distinguished as holistic or analytic. This chapter follows these distinctions between a holistic and an analytic rubric, as presented by Stanley (2021). The author explains: "Just as its name suggests, a holistic rubric looks at the whole picture. If a student is completing a performance assessment, such as an essay, the evaluator would look at all aspects of the written piece and narrow all of the criteria into a single grade" (Stanley, 2021, p. 37). As expected, these rubrics allow us as educators to look at the whole picture and score students in a faster manner. On the other hand, an analytical rubric "allows the evaluator to show the exact areas in which students display strengths and areas in which students display need for improvement" (Stanley, 2021, p. 39). Analytical rubrics, then, interact with scorers in ways in which holistic rubrics would not be able to; analytical rubrics are, therefore, all about the specifics that will help the student to recognize their faults and work thoroughly towards improvements and successes. In short, holistic rubrics should be perceived as a summative assessment as they provide information about what students can do well at a given point.

Alternatively, an analytic rubric should be perceived as a formative assessment, as these can demonstrate the areas in which students need to work looking for improvement; something particular about the analytic rubrics is that these conjointly provide feedback to educators, as they can easily observe the areas that still need to be mastered as part of a given lesson. Stanley (2021) illustrates the previous explanation as follows: "If a student writes a rough draft of the essay and the teacher uses a rubric to evaluate it, the student might learn that, although he scored an A on grammar and spelling and a B on

content, his organization was very poor, resulting in a D: (p. 39). The alternation of rubrics in the language classroom should be highly emphasized as emergent bilinguals can potentially benefit from feedback and the constant revisiting of materials, presented, of course, in varied ways, in their classrooms; by providing emergent bilinguals with a variation of rubrics, we can acknowledge equally their strengths and weaknesses creating a safe space for them for learning and embracing a new language and culture. In this way, emergent bilinguals will pose a pattern of having explicit explanations about how to perform a task appropriately by receiving a rubric in a timely manner; but at the same time, we will also give them the opportunity of learning from their different experiences with a variety of rubrics.

Additionally, Stevens and Levi (2013) take notice of the advantages of scoring rubrics by acknowledging the following: "They allow for much greater individualization and flexibility in grading" (p. 79). For this reason, when we, as educators, want to provide as much freedom as possible to our students, scoring rubrics represents a good choice. Becker (2016) highlights that "scoring rubrics appeal to teachers and students because they increase the transparency of teacher's expectations regarding the criteria for a quality performance, enhance students' awareness of learning goals, and enable the provision of feedback that can help students to identify strengths and weaknesses in their work" (p. 5). It is primarily essential for the language educator to constantly keep in mind students' learning goals since considering the amplification of their linguistic repertoires -commonly scored in many language classrooms- should not be the only learning goal to be considered; developing confidence while performing other tasks like listening or speaking should also be considered an impactful objective. In the same way, learning to self-correct and developing more complex grammatical structures should also be considered relevant learning goals when scoring them.

MAIN FOCUS: WHY DO WE NEED RUBRICS IN THE LANGUAGE CLASSROOM?

Although providing a clear route for students' understanding and success (Shabani & Panahi, 2020) is not particular to language assessment, it is unquestionably a feature that must be taken into consideration when creating and distributing them amongst students and evaluators. For this reason, using rubrics appropriately for language testing is something that educators, evaluators, and policy-makers must constantly consider when thinking about the language classroom and its assessments. Shabani and Panahi (2020) highlight essential benefits for students when presented with language assessment tools, such as rubrics; the authors mention that a rubric: "Helps learners understand raters' and teachers' expectations better, judge and revise their own work more successfully, promote self-assessment of their learning, and improve the quality of their writing task" (p.5). These benefits demonstrate in an observable way how students' linguistic practices and knowledges can improve during a specific period of time. For this reason, the present chapter emphasizes the use of rubrics compared to other assessment tools since rubrics generally may provide more objective and effective feedback for students.

Some authors have explicitly emphasized some benefits of using rubrics, or as they call it, having a marking schema. For example, authors have mentioned that rubrics:

Give confidence and guidance to novice markers and help them become part of an assessment community more quickly. They provide guidance for the experienced marker, who may have developed inaccurate marking practices. They have some standardizing effects. They provide a public language and termi-

nology for assessment requirements. They make outcomes more transparent for students, and students value them (Bloxham & Boyd, 2007).

Because emergent bilinguals heavily rely on the expectations language classroom rubrics may have (Athon, 2019), it is advisable to provide them with the best assessment tools available. The author suggests that whereas emergent bilinguals are constantly trying to replicate the categories of the rubrics that are provided to them, emergent bilinguals have the tendency to implement similar structures and patterns from the ones already pre-established in their language courses in the rest of the courses they take in the future. It is remarkable to highlight that when creating standards, and therefore when creating rubrics that are supposedly supporting such standards, language must be accessible not only to students but also to their teachers. Glass (2004) constantly reminds us that when standards or rubrics possess not an approachable language, both tend to be useless.

In spite of the fact that deciding whether or not to use an individualized rubric tends to be of choice for educators, this should come with a warning that each specific assignment, primarily in the language classroom, is addressing a specific area or skill; for this reason, the specificity a rubric may incorporate tends to be more satisfying than only providing a generic one. Novice educators, for example, could be tempted to exclusively use a generic rubric to assess all different assignments, which could quickly raise validity and reliability concerns for more advanced educators. This, consequently, should be a call for curriculum creators to provide more room for pre-service teachers to learn about the many different types of rubrics available for different fields and, equally important, to learn about their advantages when using them appropriately. Providing feedback with the use of a rubric in a language classroom can increase students' performance, as emergent bilinguals are constantly attending to feedback and looking for corrections and revisions (Ene & Upton, 2014; Ene & Kosobucki, 2016). Standard written and oral language conventions to be included in a rubric are sentence structure, grammar, punctuation and capitalization, spelling, organization, and content; nevertheless, we must adequately train pre-service teachers and novice educators to think outside the box concerning the abovementioned conventions. Currently, feedback must also include newer cultural conventions, cultural awareness, and more flexible linguistic repertoires, which might not necessarily be represented in a standard pre-made rubric.

Glass (2004) highlights that whenever we are using a pre-made rubric, it is essential to carefully review the language used to verify if it actually fits the needs of the targeted students in your own classroom. Equally important, she highlights that when creating rubrics for a program, all should have an even understanding of the language being used to assess students. Holding several discussions with interdisciplinary colleagues, the question regarding assessment in language courses always remains the same: If it is not objective, how do you assess your students? As it has been severely discussed by many (Ayhan & Türkyılmaz, 2015; Glass, 2004; Khanmohammad & Osanloo, 2009), assessing language can sometimes be objective and subjective some other times. Glass (2004) notes, for instance: "One teacher's impression of a capable job at a powerful climax might be another's advanced version: (p. 15). For this reason, the implementation of rubrics, particularly in the language classroom, should be heavily addressed. In this sense, it is advised that every language educator recurrently occurs in calibration sessions where they, along with other language educators and supervisors, can conjointly facilitate assessment practices utilizing the rubrics to be implemented in their classrooms. The author further warns the reader about the importance of balancing just the right amount of text when developing a rubric. Otherwise, rubrics can deviate students' and evaluators' impressions.

Conversely, when vaguely or up to no information is provided, students and evaluators can easily follow objectively what has been required or intended. Therefore, creating rubrics that can actually be successful for emergent bilinguals and, more importantly, which include the most appropriate conventions for them should be a task that requires time, patience, and a significant amount of collaboration; the collaborations, as mentioned earlier, should not include collaborations amongst colleagues exclusively, but no less important, should be the collaboration between educators and students.

As suggested by different authors (Glass, 2004; Stanley, 2021), educators can take a more collaborative approach when allowing their students to generate new rubrics with them. Accordingly, because the language classroom usually represents an area where minoritized students are beginning their education endeavors, it is advisable to challenge them to create essential standards collectively. In this scenario, a way to share the voices of the emergent bilinguals in these classrooms is to emphasize the importance of language mechanics and simultaneously the relevance of students' identities. Pappas, Zecker, and Zecker (2001) note: "Collaborative classroom interactions occur when teachers move away from teaching-as-transmission approaches to ones in which they share power and authority with their students" (p. 7). This should be highly considered in the language classroom, where many students may sometimes feel powerless given their social, cultural, and linguistic experiences. Additionally, guiding our emergent bilinguals to learn how to collaborate to achieve success can help them remediate the above mentioned circumstances.

One possibility when taking this approach of rubric-creation with our own students in the language classroom would be to clearly model sentence patterns to be used in order to create the expected rubric. Glass (2004) recommends: "When you take this approach, though, have a keen sense of what you expect in the writing assignment, and prod the students while they brainstorm criteria to make sure all essential elements are included" (p. 23). Following a student-centered learning approach, like guiding emergent bilinguals during their brainstorming sessions, can enhance students' willingness to practice the target language voluntarily and simultaneously can also provide insights for us as language educators better to support struggling emergent bilinguals (Unin, 2016). Stanley (2021) highlights how incorporating non-traditional activities in the language classroom, such as the co-creation of materials and rubrics, can serve as a motivational tool. The author states: "Rubrics can be used to teach students to track their own growth, making them self-reflective learners. They could have the rubric ahead of time and then be able to take their work and progress and assess their level of understanding and mastery" (p. 32). For this reason, it is highly suggested that educators take different routes when trying to empower their students. For emergent bilinguals, helping them to become reflective learners can have a tremendous positive impact as they may feel ashamed of continuously having to ask the same questions, or very similar questions, in an effort to follow what a lesson might require from them entirely.

Authors (Anthon, 2019; Stanley, 2021) have highlighted how positively impactful it is for educators to allow students to create or co-create their own rubrics. Consequently, letting students help educators in the development of their own learning rubrics is beneficial for students learning processes. Stevens and Levi (2013) note:

Integrating rubric construction into classroom teaching can: (1) prevent misunderstandings and misinterpretations before they affect student work; (2) increase student awareness of themselves as 'stakeholders' in the educational process, which, in turn, results in greater student involvement in the tasks assigned; (3) cut down your workload by letting your students do some of it (Stevens & Levi, 2013, p. 49).

By following such an approach, emergent bilinguals are taught to collaborate responsibly. In this sense, emergent bilinguals are offered opportunities to elaborate and reflect on their diversified language learning processes. Meanwhile, by collaborating with others, educators or classmates, emergent bilinguals are able to integrate different elements in their language learning which can be later displayed, for instance, in a written or spoken format (Bradley, Lindström, & Rystedt, 2010). Interestingly, we are not only supporting approaches that are closely related to the experiences emergent bilinguals will be having outside of the classroom by providing authentic collaboration practices, but conjointly, we are also lightening our workload by creating more responsible students.

As Bloxham and Boyd (2007) observe, "Feedback is arguably the most important aspect of the assessment process in raising achievement" (p. 103). Consequently, by using rubrics constructively to provide feedback, optimistically timely feedback, educators can encourage positive motivation amongst students and, at the same time, can help students move forward in their learning processes as they count with effective written feedback which can strengthen their opportunities for improvement. Especially in language classrooms, the possibilities for assessments may vary drastically depending not only on the level of proficiency of the emergent bilinguals, but also their level of education, level of engagement, and even their cultural backgrounds. For this reason, the constant use of rubrics in a language classroom may help emergent bilinguals to constantly mobilize their diversified knowledges to continue scaffolding their opportunities for learning and success.

Validity and Reliability

Authors have noted (Bloxham & Boyd, 2007; Stevens & Levi, 2013) the reliability and validity of the rubrics used in a classroom are continuously impacted not only by the assessment criteria stated but by the actual marker. For this reason, it is advisable to work collaboratively to practice how students are to be assessed and, conjointly, how rubrics are to be interpreted and used to assess them. Stevens and Levi (2013) note that rubrics constantly change. Because of these changes mentioned above, rubrics are also becoming better and more well-adjusted to the needs of a given classroom. Because of the many specificities that a language classroom requires, working effectively towards the constant creation and modification of rubrics that can be easily adapted to the needs of emergent bilinguals will continue supporting the improvement, validity, and reliability of such rubrics. Bloxham and Boyd (2007) address how important it is to provide rubrics that not only validly assess students but, more importantly, transparently are designed to match their learning outcomes as their programs and institutions have observed these. Like highlighted by other authors (Athon, 2019; Glass, 2004), presenting information in rubrics using a language that the students do not welcome, may diverge their interest in reading the rubric leading to obtaining a less satisfactory accomplishment.

Jeong (2015) notes that although the use of rubrics, specifically for language assessment, has increased potentially during the last years, the question remains as to up to what point the users/evaluators are true to their meaning; in other words, how they accurately use rubrics in comparison to other users/evaluators in their same position. Jeong (2015) reminds us of the relevance of using rubrics appropriately because students may be concerned about how they have been graded. Therefore, just like students' knowledges must be assessed accurately and reliably, at the same time, rubrics must also be accurate and reliable. Some other authors (Jeong, 2015; Turley & Gallagher, 2008; Wilson, 2007) also highlight the importance of using the rubrics appropriately and having appropriate training to use them efficiently. Although the present chapter focuses on the relevance of using rubrics for language assessment, it would be advisable

to consider using a rubric not in isolation but in conjunction with other assessment tools to exalt their validity and reliability.

In order to provide a more accurate, and thus objective, assessment in the language classroom, Glass (2004) recommends the following: (1) To create and revise rubrics; (2) to score student work with colleagues; (3) to collect student anchor papers; (4) to avoid bias; and (5) to use student scores to inform instruction. It is relevant for educators and curriculum developers to incorporate the appropriate use of rubrics in the language classroom. Most importantly, they have to ensure that such rubrics are used without bias and closely follow students' actual work. By considering the collection of students' samples, we are not only empowering their learning processes, but concertedly, we are also practicing scoring assignments, and their corresponding rubrics, that are authentic and of a higher value for our own programs. On the one hand, we are empowering students as we are providing them with the opportunity of becoming, for instance, the ideal writer. On the other hand, we are collectively potentializing our opportunities as language educators to clearly provide observable samples of the strengths or weaknesses the lessons we are giving are producing in our students. By keeping these samples and scoring them collectively constantly, we will soon feel more comfortable navigating the rubrics that are required to assess our students in the language classroom properly.

Language-Specific-Skills and Rubrics

Even though using rubrics in the language classroom has been becoming a quite often practice among language educators, finding rubrics that are especially targeted to assess a specific language skill is not as easy; for the most part, research and resources are developed to gain a better understanding of the writing learning processes of emergent bilinguals. Nevertheless, other skills such as listening, reading or speaking are relegated. Assessing writing proficiency, as discussed previously, has been growing significantly in the last years, as emergent bilinguals are expected to learn how to appropriately undergo specific writing conventions, particularly in English. Shabani and Panahi (2020) argue that: "Learners are generally expected to produce a piece of text so that raters can evaluate the overall quality of their performance using a variety of different scoring systems including holistic and analytic scoring, which are the most common and acceptable ways of assessing essays" (p. 3). Because writing is considered as the most crucial skill in language learning, and also in a variety of other fields, the implementation of the use of rubrics is a well-established practice which is constantly looking for quality in all its respects.

Furthermore, considering that writing is a self-regulated learning process (Hawthorne & Pribesh, 2017), the constant implementation of rubrics to provide detailed feedback for emergent bilinguals should be a constant in the language classroom as a means to help in students' writing development. Hawthorne, Bol, and Pribesh (2017) note: "Well-developed, intentionally designed rubrics provide students with evaluation criteria and the opportunity to think about their writing and to compare and evaluate their success against identified standards" (p. 4). Consequently, by providing enough encounters in our language classrooms where students are exposed to the use of rubrics we can help them apply more effective writing strategies that will conjointly help them to move forward in their writing learning processes. Additionally, by providing emergent bilinguals with rubrics that clearly represent what is highly valued by many institutions (e.g., correct spelling), we can help students in our language classroom to learn how to better self-regulate the skills that are required to become a successful writer (Andrade, 2000; Saddler & Andrade, 2004).

Even though speaking is not as commonly assessed as writing, it has becoming more evident that its relevance in the process of language development plays a crucial role; for this reason, finding rubrics that help us as language educators to evaluate better and provide timely, consistent, and relevant feedback to emergent bilinguals must be kept in mind. In regards to speaking rubrics, Pineda (2014) notes: "Performance assessments contribute to measuring students' abilities to respond to real-language tasks, value students' true language abilities, and reflect on how students will perform in future real-life language situations" (p. 186). However, using rubrics to assess speaking tends to be more time-consuming in comparison to assessing other language skills. In order to convey a truly understandable meaning when speaking, emergent bilinguals must be able to understand and provide a social and contextual appropriate meaning, possess phonological and phonetic awareness, and follow the commonalities of speech; assessing speaking through the use of rubrics can provide a straightforward way for educators and emergent bilinguals to fully understand the speaking objectives to be accomplished in a given assignment or course, and at the same time, can also help to evaluate students' vocal varieties, ways to articulate, preferred languaging practices, and nonverbal cues (Schreiber, Paul, & Shibley, 2012). Rubrics assessing speaking then help us as language educators to become better at assessing this skill by providing us, and emergent bilinguals, with opportunities to gain understanding of the messages they are creating when presenting information through their oral repertoires. The authors finalize their thoughts by suggesting that "the pedagogical and institutional advantages of descriptive rubrics suggests that they may be better candidates for providing a reliable way to assess public speaking proficiency for both expert and nonexpert evaluators" (p. 212). As it has been evaluated throughout the chapter, certain types of rubrics come more handy when trying to assess certain skills, assignments, or courses over others.

Reading as a receptive skill, however, could be more well-represented when looking for evidence about the efficiency of rubrics. This does not come as a surprise since reading is a mental activity which, therefore, must be evaluated through other language skills like speaking or writing (Blaz, 2001). Some observable behaviors that can be identified easily when trying to assess reading in a language classroom could be for example underlying unfamiliar vocabulary and providing definitions or translations for such unfamiliar words; when working with unfamiliar vocabulary in the language classroom, we should avoid preventing students from only using the target language since language practices are to be constantly mobilized and expanded instead of remain static (Garcia & Wei, 2015). When assessing reading, rubrics in the language classroom can, for example, illustrate the level of achievement an emergent bilingual may have in terms of reading comprehension (Leist, Woolwine, & Bays, 2012). In addition, reading rubrics can help us as language educators to better evaluate more complex reading tasks, like their progression when being able to evaluate, critique, or predict while reading (Mijušković, 2014). Other aspects that can be also beneficial when assessing reading through the use of rubrics are phonological awareness, recognition of lexical patterns, ability to make inferences and connections, and last but not least, ability to provide background information. Because reading is a scaffolding skill to be developed constantly, it is recommended to use varied rubrics but in a constant manner, that way reading skills can be evaluated from an array of perspectives.

Unfortunately, the scarcity of research in relation to cultural bias and its direct relation to assessment (Bloxham & Boyd, 2007) may still put us in a difficult position as language educators. The authors highlight that "there is evidence that different types of assignment or examination differentially impact on international students' ability to demonstrate their learning" (p. 152). Therefore, in order to promote a more inclusive and equitable type of rubric in the language classroom, educators, curriculum creators, and more importantly, assessment creators, must provide clear expectations for all given assignments;

similarly, they should consider that ethnic differences may also impact students' understanding of some other concepts such as plagiarism, leading to uninformed practices when completing an assignment. It is then important to finely discuss with students any language used in a given rubric that might be disruptive, vague, or biased. As previously mentioned, when providing standards and rubrics for students, language must be as accessible and clear as possible; otherwise, we are not only accumulating checklists that are unattainable and uninformative. To diminish cultural-bias, Kim and Zabelina (2015) suggest staying away from standardized testing and consider instead other alternative assessments which give more room to creativity. The authors highlight how even those students who might look acculturated may also suffer from cultural bias. Consequently, language educators should be really self-critical when deciding how to assess students to avoid any cultural bias.

Based on Inoue (2015), Athon (2019, p. 79) states:

Inoue discusses his use of grading contracts in his first-year composition courses where the letter grade values labor over any single writing product. Students, regardless of writing ability, must regularly attend class and complete all assignments to earn a high grade. By acknowledging the labor of writing as the act of learning—rather than a single essay privileging a white language variety—Inoue argues that we can lessen assimilation to a dominant discourse (Athon, 2019, p. 79).

Making decisions about what should have a heavier weight when assessing emergent bilinguals can represent a challenge. Following Inoue's (2015) observations, language educators, as well as policy makers, should carefully address, and more importantly acknowledge, the many steps involved in learning for an emergent bilingual. For this reason, we, as language educators, should purposefully consider many alternatives in our own classrooms that instead of continuing to privilege dominant discourses, can provide more supportive practices for our students. Added to this, providing students with clear expectations by the constant provision of rubrics will leverage their learning processes as students can learn to appreciate the value of such rubric; the constant dissemination of rubrics in the language classroom which clearly replicates the labor they are constantly involved in, and not necessary just the product, will provide emergent bilinguals with additional tools that will guide them to freely show their agency teaching them how capable of learning they already are.

Stanley (2021) clearly states many advantages that using rubrics specifically created by the educator to assess a particular class may entail; for instance, because educators know students' proficiency levels and needs, they can customize what could be more beneficial for their own students. In this sense, by promoting that educators create their own rubrics, they can better help their own students to scaffold their learning processes. This applies particularly to the language classroom where students are constantly challenged by many other factors, in addition to the content. Stanley (2021) pays particular attention to the use of rubrics to assess authentic experiences, which hopefully in today's language classroom are constantly occurring. The author continues "Advanced students need to be challenged more, while struggling students might need more scaffolding to succeed. Rubrics are naturally scaffolded. Because there are different levels of performance, if you have a class with a wide spectrum of student ability, you could use the rubric as a way to differentiate" (p. 29). With that, assessing emergent bilinguals by using rubrics accurately can diminish having students who are not only poorly supported academically, but who could be more educationally successful.

SOLUTIONS AND RECOMMENDATIONS

It is unquestionable that rubrics will only be useful for students if educators know how to work with such rubrics. For emergent bilinguals particularly, special attention is to be paid to the practices (linguistic, contextual, and social) that are addressed, privileged, incorporated, and taught, when providing them with a rubric. Consequently, the chapter aims to make a call to curriculum developers and creators of faculty development to continue providing opportunities to educators and future educators to learn deeply about the usages and benefits of rubrics in the language classroom. It is worth reminding the reader that, many times, rubrics are used in isolation when assessing language; nevertheless, conjoint assessing tools could help students' learning processes.

Many authors (Athon, 2019; Becker, 2016; Panadero & Jonsson, 2013; Reddy & Andrade, 2010; Stevens & Levi, 2013; Wolf & Stevens, 2007) address the importance of involving emergent bilinguals in the creation and co-creation of rubrics for the language classroom. Becker (2016) states: "While students are often introduced to a scoring rubric prior to or at the end of a writing assignment, they are rarely involved in the development and application of scoring rubrics" (p. 4). Nevertheless, the advantages of collaborating with our own students to create and co-create rubrics are numerous. For example, by co-creating rubrics with our students, they may be able to internalize their knowledge better as they become self-aware of what they are expected to learn (Fraile, Panadero, & Pardo, 2017). In the same way, co-creating rubrics with our students may enhance the empathy educators are able to establish with students, but most importantly, this practice may create a sense of ownership from the students' side (Joseph, Rickett, Northcote, & Christian, 2020) learning how to promote their own voice (Fraile, Panadero, & Pardo, 2017). In this sense, the implementation of this strategy would alternatively provide educators with an assessment tool which is fully understandable and transparent for the students, and additionally, which will lead students to use their own voice.

FUTURE RESEARCH DIRECTIONS

One evident disadvantage of the current research on rubrics for language assessment is the limited literature that can be found on assessing language skills other than writing. It would be worth it, however, to continue researching all language skills equally to provide more options for educators while assessing. Another possible contribution could be to research the benefits of using rubrics compared to other assessing tools. Equally important, research contributing to the dissemination and experiences of educators and students as they collaboratively co-create rubrics would be of value. Researching the impact of digital rubrics versus paper-based rubrics could also be beneficial when making choices in our language classrooms.

CONCLUSION

Although the raters' knowledge when using a rubric appropriately has not, to the author's knowledge, been widely researched in comparison to other elements, research particularly to language assessment has highly contributed to the features that should be tested in language; these features are mostly relevant for writing, but can be used to assess other language components; Shabani and Panahi (2020) included

the following elements for assessment: Content, organization, cohesion and coherence, vocabulary and grammar, and language and mechanics (p. 7). As previously mentioned, although these are mostly valued and used to assess writing, other language abilities can also be assessed in a similar fashion.

As it has been revisited during the chapter, the level of specificity provided as part of a given rubric is commonly mandated by the educator or the curriculum creator. In this sense, although having a general rubric can serve various purposes in a class, it should be evident by now that this will not be as beneficial as providing a less general rubric. On the other hand, having task-based rubrics provide more specifics, and therefore, allow students to become familiar with the expectations promptly. Furthermore, task-based rubrics provide a more ample panorama when trying to track back what was not accomplished when completing a task; this way, educators can reinforce what did not go that well in a lesson, but conjointly, students can also become more responsible for their own learning by becoming easily aware of what they can still improve as part of their learning process. We must remember at any point that avoiding cultural bias, for instance when providing rubrics to students, will promote a more inclusive environment for our emergent bilinguals.

Furthermore, Brookhart and Chen (2015, p.363) remind us: "teachers or post-secondary faculty should have professional development in using rubrics and in coaching students to use rubrics, and pre-service teachers should have training in these matters, as well". For this reason, this chapter intends to further the understanding of the value of rubrics in the classroom, more specifically, the value of a rubric in a language classroom. More importantly, the chapter looks to provide evidence of the current need for educators, language educators for the most part, to be appropriately trained to create and use rubrics in their classrooms. For this reason, the current chapter not only provides general distinctions among the different types of rubrics that we have available as educators, but also, it provides specifics in relation to rubrics directly assessing the core language skills (reading, writing, speaking, and listening). As a concluding thought, the present chapter looked to contribute to gaining a deeper understanding of some of the most common types of rubrics used for language assessment as well as their benefits when using them appropriately.

REFERENCES

Andrade, H. G. (2000). Using rubrics to promote thinking and learning. *Educational Leadership*, *57*(5), 13–19.

Athon, A. (2019). Designing Rubrics to Foster Students' Diverse Language Backgrounds. *Journal of Basic Writing*, *38*(1), 78–104. doi:10.37514/JBW-J.2019.38.1.05

Ayhan, Ü., & Türkyılmaz, M. U. (2015). Key of language assessment: Rubrics and rubric design. *International Journal of Language and Linguistics*, *2*(2), 82–92.

Becker, A. (2016). Student-generated scoring rubrics: Examining their formative value for improving ESL students' writing performance. *Assessing Writing*, *29*, 15–24. doi:10.1016/j.asw.2016.05.002

Blaz, D. (2001). *A collection of performance tasks and rubrics: Foreign languages*. Eye On Education.

Bloxham, S., & Boyd, P. (2007). *Developing effective assessment in higher education: a practical guide: a practical guide*. McGraw-Hill Education.

Bradley, L., Lindström, B., & Rystedt, H. (2010). Rationalities of collaboration for language learning in a wiki. *ReCALL*, *22*(2), 247–265. doi:10.1017/S0958344010000108

Brookhart, S. M., & Chen, F. (2015). The quality and effectiveness of descriptive rubrics. *Educational Review*, *67*(3), 343–368. doi:10.1080/00131911.2014.929565

Ecclestone, K. (2001). 'I know a 2: 1 when I see it': Understanding criteria for degree classifications in franchised university programmes. *Journal of Further and Higher Education*, *25*(3), 301–313. doi:10.1080/03098770126527

Ene, E., & Kosobucki, V. (2016). Rubrics and corrective feedback in ESL writing: A longitudinal case study of an L2 writer. *Assessing Writing*, *30*, 3–20. doi:10.1016/j.asw.2016.06.003

Ene, E., & Upton, T. (2014). Learner uptake of teacher electronic feedback in ESL composition. *System*, *46*, 80–95. doi:10.1016/j.system.2014.07.011

Fraile, J., Panadero, E., & Pardo, R. (2017). Co-creating rubrics: The effects on self-regulated learning, self-efficacy and performance of establishing assessment criteria with students. *Studies in Educational Evaluation*, *53*, 69–76. doi:10.1016/j.stueduc.2017.03.003

García, O., & Wei, L. (2015). Translanguaging, bilingualism, and bilingual education. The handbook of bilingual and multilingual education, 223-240.

Glass, K. T. (2004). *Curriculum design for writing instruction: Creating standards-based lesson plans and rubrics*. Corwin Press.

Hawthorne, K. A., Bol, L., & Pribesh, S. (2017). Can providing rubrics for writing tasks improve developing writers' calibration accuracy? *Journal of Experimental Education*, *85*(4), 689–708. doi:10.1080/00220973.2017.1299081

Inoue, A. B. (2015). *Antiracist writing assessment ecologies: Teaching and assessing writing for a socially just future*. Parlor Press LLC. doi:10.37514/PER-B.2015.0698

Jeong, H. (2015). Rubrics in the classroom: Do teachers really follow them? *Language Testing in Asia*, *5*(1), 1–14. doi:10.118640468-015-0013-5

Joseph, S., Rickett, C., Northcote, M., & Christian, B. J. (2020). 'Who are you to judge my writing?': Student collaboration in the co-construction of assessment rubrics. *New Writing*, *17*(1), 31–49. doi:10.1080/14790726.2019.1566368

Khanmohammad, H., & Osanloo, M. (2009). *Moving toward objective scoring: A rubric for translation assessment*.

Kim, K. H., & Zabelina, D. (2015). Cultural bias in assessment: Can creativity assessment help? *The International Journal of Critical Pedagogy*, *6*(2).

Lee, Y. W., Gentile, C., & Kantor, R. (2010). Toward automated multi-trait scoring of essays: Investigating links among holistic, analytic, and text feature scores. *Applied Linguistics*, *31*(3), 391–417. doi:10.1093/applin/amp040

Leist, C. W., Woolwine, M. A., & Bays, C. L. (2012). The effects of using a critical thinking scoring rubric to assess undergraduate students' reading skills. *Journal of College Reading and Learning*, *43*(1), 31–58. doi:10.1080/10790195.2012.10850361

Mijušković, M. (2014). Assessing students' reading comprehension through rubrics. *Mediterranean Journal of Social Sciences*, *5*(13), 252.

Panadero, E., & Jonsson, A. (2013). The use of scoring rubrics for formative assessment purposes revisited. *Educational Research Review*, *9*, 129–144. doi:10.1016/j.edurev.2013.01.002

Pappas, C. C., Zecker, L. B., & Zecker, L. (2001). *Teacher inquiries in literacy teaching-learning: Learning to collaborate in elementary urban classrooms*. Routledge. doi:10.4324/9781410600769

Pineda, D. (2014). The feasibility of assessing teenagers' oral English language performance with a rubric. *Profile Issues in TeachersProfessional Development*, *16*(1), 181–198. doi:10.15446/profile.v16n1.43203

Quinlan, A. M. (2012). *A complete guide to rubrics: Assessment made easy for teachers of K-college*. R&L Education.

Reddy, Y. M., & Andrade, H. (2010). A review of rubric use in higher education. *Assessment & Evaluation in Higher Education*, *35*(4), 435–448. doi:10.1080/02602930902862859

Saddler, B., & Andrade, H. (2004). The writing rubric. *Educational Leadership*, *62*(2), 48–52.

Schreiber, L. M., Paul, G. D., & Shibley, L. R. (2012). The development and test of the public speaking competence rubric. *Communication Education*, *61*(3), 205–233. doi:10.1080/03634523.2012.670709

Shabani, E. A., & Panahi, J. (2020). Examining consistency among different rubrics for assessing writing. *Language Testing in Asia*, *10*(1), 1–25. doi:10.118640468-020-00111-4

Stanley, T. (2021). *Using rubrics for performance-based assessment: A practical guide to evaluating student work*. Routledge. doi:10.4324/9781003239390

Stevens, D. D., & Levi, A. J. (2013). *Introduction to rubrics: An assessment tool to save grading time, convey effective feedback, and promote student learning*. Stylus Publishing, LLC.

Turley, E. D., & Gallagher, C. W. (2008). On the" uses" of rubrics: Reframing the great rubric debate. *English Journal*, 87–92.

Unin, N., & Bearing, P. (2016). Brainstorming as a Way to Approach Student-centered Learning in the ESL Classroom. *Procedia: Social and Behavioral Sciences*, *224*, 605–612. doi:10.1016/j.sbspro.2016.05.450

Wilson, M. (2007). Why I won't be using rubrics to respond to students' writing. *English Journal*, *96*(4), 62–66. doi:10.2307/30047167

Wolf, K., & Stevens, E. (2007). The role of rubrics in advancing and assessing student learning. *The Journal of Effective Teaching*, *7*(1), 3–14.

KEY TERMS AND DEFINITIONS

Core Language Skills: The current chapter constantly refers to the core language skills when presenting certain benefits of the use of rubrics. The core language skills included here are reading, writing, speaking, and listening; nevertheless, the understanding of culture and cultural backgrounds are also constantly addressed as part of the aforementioned core language skills.

Dominant Discourses: As part of the present chapter, dominant discourses include primarily the use of Academic English, and/or Standard English. These discourses can also refer to the use of Academic Languages or Standard Languages.

Emergent Bilinguals: For the purposes of the present chapter, the term emergent bilinguals is preferred over other well-known terms, such as language learner, as the first one provides a more inclusive way of presenting students. Emergent bilinguals are to be understood as those students who are in the constant process of acquiring one or more languages.

Evaluators: People providing an evaluation when using a rubric. The chapter uses the terms scorer and evaluator interchangeably. In this context tend to be educators or Teaching Assistants.

Rubric: A rubric in the present chapter is understood as an assessment tool, primarily targeted to the assessment of language and language conventions.

Scorers: People providing a score to a given rubric; people using a rubric to score students. Scorers in this context tend to be educators or Teaching Assistants.

Task-Based Rubrics: Rubrics that are looking specifically to assess certain tasks (in the language classroom for the purposes of the chapter).

Chapter 2
The Use of Rubrics for Drawing Graphs in Physics Education

Işıl Aykutlu

https://orcid.org/0000-0003-4068-0453

Hacettepe University, Turkey

ABSTRACT

In physics classes, graphs are utilized for various topics to make sure students better understand the relationship between variables. Graphs are especially useful for showing the relationship between dependent and independent variables in experiments carried out in lab classes. There are various sub-steps to drawing a graph. In order to draw a graph, students should have a good grasp of physics subjects; moreover, they should know the characteristics of graphs and be able to form the graph according to the steps. Using rubrics to assess students' drawings at the end of this process, which involves multiple skills, would ensure that the assessment is both more detailed and more reliable and valid. Moreover, using rubrics in teaching how to draw a graph would contribute to students' understanding of the graphs' characteristics; it would also help students clearly see the steps as well as contribute to the improvement of students' drawings through self and peer assessment.

INTRODUCTION

Some subjects within physics include abstract concepts. This makes it difficult for students to understand the subjects and decreases their level of success. Visual materials are used to concretize abstract concepts so that they could be more easily understood. Using visual materials such as figures, schemas, graphs, and pictures enables more effective teaching by increasing students' curiosity and motivation (Arpaguş et al., 2011). Graphs, one of these visual materials, make data easy to understand, summarize, arrange, interpret, and present; they can provide information in an effective way various data sets cannot (Demirci & Uyanık, 2009; Taşar et al., 2002). Line graphs are frequently used especially in physics classes to show the relationship between concepts. Visually presenting the relationship between two continuous dependent and independent variables, line graphs are an indispensable part of experiments in positive sciences (McKenzie & Padilla, 1986). The relationship between various concepts in physics classes is

shown by employing graphs. In order for students to understand the relationship between these physics concepts, they need to understand and interpret graphs correctly (Demirci & Uyanık, 2007). In this respect, their graph skills or their graph reading-interpretation and drawing skills, should be at an adequate level. Studies on physics education show that students' graph drawing skills are lower compared to their graph reading-interpretation skills (Eryılmaz-Toksoy, 2020; Yeltekin Atar & Aykutlu, 2023). Drawing a graph involves the following consecutive steps: selecting an axis, tagging an axis, scaling the axis, data entry, forming a point, and connecting the points (Aydın &Tarakçı, 2018; Gültekin 2009; McKenzie & Padilla, 1986; Yeltekin Atar & Aykutlu, 2023). Students' drawings are influenced by different skills and knowledge such as their knowledge of physics, mathematics, and rationalization. This complex and multifaceted nature of graph drawing makes it challenging to determine the difficulties students face and to improve their skills (Angra & Gardner, 2018). In this respect, students' graph drawings should not be assessed as correct or incorrect. Studies show that rubrics can be used to teach students how to draw graphs and then to assess them (Angra & Gardner, 2018; Yeltekin Atar & Aykutlu, 2023). By using rubrics designed according to pre-determined criteria, students can be shown to comprehend what is expected of them when it comes to drawing a graph related to physics topics. Moreover, students' performances can be assessed in detail by using rubrics to evaluate their graph drawings. In light of all this information, this study aimed to assess how rubrics could be employed in evaluating students' drawings and in teaching them how to draw graphs. It is believed that examining how rubrics could be used in drawing graphs in physics education would be significantly beneficial for physics teachers, field experts, and students in physics classes. Consisting of six parts, namely, Introduction, Method, Background, The Use of Rubrics for Drawing Graphs, The Advantages and Disadvantages of Using Rubrics for Drawing Graphs, and Conclusion, this study is believed to contribute greatly to different disciplines such as mathematics, positive sciences, and social sciences, in which students utilize graphs.

METHOD

Having examined studies in the literature on rubrics, this study was designed as a traditional review. Traditional/literature/narrative/descriptive reviews examine two or more studies on a specific subject, and they are carried out to introduce a specific topic, to summarize it, or to fill a gap within the literature (Karaçam, 2013; Yılmaz, 2021). Traditional review studies, which are carried out by experts in the field, can be sometimes defined as "compilation, review, theoretical analysis, literature review" (Yılmaz, 2021, p. 1460). In this traditional review, rubrics, and the studies on the use of rubrics in physics education and graph drawing, were examined, and the ways rubrics can be used for drawing graphs were inspected.

BACKGROUND

In this part of the study, general information in the relevant literature concerning rubrics, as well as studies which use rubrics for different ends in physics education were presented.

Used in evaluating performance, rubrics consist of pre-determined performance criteria to assess students' work (Mertler, 2001). Rubrics are not only a tool to assess students, but they also allow students to see their own performance (Andrea, 2005; Oaklef, 2009). Providing feedback on to what extent students have reached the pre-determined criteria and how they could improve their performance (Moskal, 2000),

rubrics can contribute to the improvement of the quality of teaching, in addition to evaluating students' work when they are designed properly (Popham, 1997). Rubrics have three basic aspects, namely, assessment criteria, quality definitions, and scoring strategy. Assessment criteria means criteria used to distinguish students' acceptable and unacceptable answers. Quality definitions are used to determine how to assess the qualitative differences in students' answers. Quality definitions include a detailed explanation on what the student's answer should be in order to be considered perfect or weak. Scoring strategy for a rubric could be holistic or analytical (Popham, 1997; 2007). Whether it will be holistic or analytical is decided based on the aim of the assessment (Moskal, 2000). When the assessment is generally based on students' answers or on the result of their work, a holistic scoring is employed. In the analytic scoring strategy, a student's answer or work is scored step by step. Analytical rubrics focus on all of the steps effective in revealing the student's performance. In holistic rubrics, on the other hand, the student's work is assessed as a whole (Kan, 2009).

Rubrics are valuable assessment tools for both the teachers and students because they can reveal a student's performance in class quickly and in a transparent manner (Hafner & Hafner, 2003). While students feel less anxious because they would know what is expected of them and how their performance will be assessed, teachers can also use rubrics in planning a more effective teaching and in determining learning outcomes (Andrade & Du, 2005). A well-designed rubric can also be used to monitor students' improvement in relation to learning outcomes (Arter, 2002). As assessment tools, rubrics can also be used as a teaching tool because they indicate what is expected of students in terms of their performance levels (Aktaş & Alıcı, 2018). Students can see what they have learned so far and what else they should learn and assess their performance by using rubrics for self-assessment (Oakleaf, 2009). Guiding students to improve and complete their studies, rubrics can also be used between teachers, students, and parents as a communication tool about students' strengths and weaknesses. Rubrics make it easy for the teachers and students to understand the expected performance (Hal & Salmon, 2003). In addition to improving teaching, rubrics provide feedback for students, and they contribute to a fair and accurate assessment. They are also important sources of information for developing programs (Wolf & Steven, 2007).

Clarifying teachers' expectations, rubrics show students how they could meet these expectations (Andreda, 2000). Rubrics makes it possible to have self-assessment and peer assessment, and help students identify and solve issues in their studies (Andrea, 1997; Wolf & Steven, 2007). Because rubrics clearly lay out the expected criteria and provide self-assessment and feedback, they promote learning and improve teaching (Jonsson & Svingby, 2007). They are easy to use and explain, and help teachers assess students' work more quickly (Andreda, 1997; 2000). When rubrics are used to identify what to do and what points to take into consideration during a performance, students' skills and perceptions could be improved (Güneş & Kılıç, 2016). Supporting students' learning and complex thinking abilities, rubrics can also be educative if prepared with students (Andrade, 2000). Evidently, rubrics could be used to plan and improve teaching, to assess student performance and identify points to improve, and for students to see their mistakes through self- and peer assessment.

The Use of Rubrics in Physics Education

Studies on the use of rubrics in physics education show that they could be used in various forms and for numerous objectives, such as improving and assessing students' problem solving performances (Chasteen et al., 2012; Doktor et al., 2016; Teodorescu et al., 2014), improving their report writing in lab studies (Faletič & Planinšič, 2020), and examining programs (Chasteen & Scherr, 2020) text books (Munezero

et al., 2022; Papakonstantinou & Skoumios, 2021) and students' graph drawings (Bahtaji, 2020; Nixon et al., 2016; Yeltekin Atar & Aykutlu, 2023).

To contribute to students' acquisition of scientific process skills and to help them self-assess, Etkina et al. (2006) used rubrics and formative assessment tasks in their studies. This study aimed to ensure students acquire skills through activities. At the end of the study, it was determined that students' skills of designing an experiment, of designing a mathematical procedure for the solution of an experimental problem, and of evaluating experimental uncertainties had improved. Moreover, the study also found out that students could transfer these newly acquired skills to new contexts. Another conclusion of this study was that assessment tasks significantly increased students' problem-solving success.

Aiming to design a rubric to solve questions about Newton's Law of Motion and to show the efficacy of this rubric in solving problems, Kocakülah and Aytaç (2007) carried out their study with the participation of second year university students. As a first step in the study, students, who had completed the unit on Newton's Law of Motion, were asked eight questions. Then, students were informed about the aims, types, and designs of rubrics; they were asked to design their own rubrics to solve the questions on Newton's Law of Motion. After consulting with experts, the rubric to be used in evaluating students' problem-solving skills was decided. At the end of the study, it was determined that using rubrics increased students' success in problem-solving. In another study, Kocakülah (2010) examined students' performance in relation to Newton's Laws of Motion. In this study, the necessity of drawing free body diagrams, the consistency between the participants' scaling and the expert coders' scaling, and the relationship between the coders' gender and the participant scores were examined. Using a semi-experimental pattern with a pre-test – post-test control group, the participants were asked four questions before and after the application. Participants in the experiment group were informed about the use of rubrics and their characteristics, and they were asked to design their own rubrics. Participants in the control group were neither informed about nor asked to design rubrics. Participants in the experiment group were divided into groups in which they designed their rubrics; these rubrics were presented in class and were assessed through discussions. Scoring of the presentations were done independently by the participants, the researcher, and experts. The designed rubric was given to the participants in the experiment group so that they could use it when solving problems on Newton's Laws of Motion; the participants were informed that their solutions would be assessed by the researcher and their peers. At the end of the study, it was seen that the designed rubric increased students' success and also provided consistency in the scoring.

In their study, Hull et al. (2013) analyzed students' problem-solving strategies by using rubrics. Problem solutions of two students of the same kinematic problem were presented in the study. It was seen that even though one solution contained more expertise compared to the other solution, the scores these students received after the use of a rubric were approximately the same. At the end of the study, it was also determined that the rubrics were unable to differentiate in the analysis of students' solutions. To assess students' practical physics skills, Liew et al. (2019) designed a scoring rubric in their study. This rubric is valid and reliable to assess student's practical physics skills in secondary education physics labs. In his study carried out to improve students' graph drawing skills and conceptual understanding, Bahtaji (2020) used rubrics to assess students' graphs. The designed rubric was used in scoring and classifying the graphs drawn by students.

In their study, Faletič and Planinšič (2020) examined the efficacy of using self-assessment rubrics when preparing student reports on open-ended experimental physics questions. The rubric used by students when preparing their reports was the same rubric used by the educators when scoring these

reports. Students' reports were examined by the educator and sent back to the student so that they could improve it. This process continued until the report was accepted. At the end of the study, it was seen that the use of a rubric decreased the burden of the scorer. Moreover, the use of a rubric also increased the quality of students' reports. Explaining to students what rubrics are and how they are used can contribute to minimizing the experimental uncertainties in lab studies. Chasteen and Scherr (2020) developed a rubric for measuring the characteristics of a successful physics education program. In addition to defining the existing applications and structures in the physics education program, this rubric also contained an objective and reliable self-assessment. Moreover, it was determined in the study that rubrics could be used to reveal the areas of improvement to consider strengthening the program.

To assess engineering students' problem-solving processes in physics and mathematics, Salazar-Torres et al. (2021) developed an analytical rubric. This rubric consists of four criteria containing Polya's problem solving steps. These criteria are "interpreting the problem," "configure a plan," "execute the plan," and "look back." At the end of the study, four analytical rubrics were formed for each criterion. It was stated in the study that analytical rubrics provide an easy-to-understand and fair assessment for both teachers and students when assessing the solution of physical and mathematical problems. Kocakülah (2022) developed and used a rubric to assess university students' solution suggestions for questions on electromagnetic induction. The study used a semi-experimental pattern with a pre-test – post-test control group, and the students in the experimental group were asked to develop a rubric to assess their problem solutions for electromagnetic induction. At the end of the study, it was determined that the developed rubric was helpful for identifying the characteristics of the problem solutions and increasing student success. Moreover, the rubric was also seen to measure students' scores consistently, independent of the scorer. Evidently, the rubric could be used in the teaching of electromagnetic induction to deepen students' conceptual understanding.

All these studies show that the use of rubrics increases student success, helps identify students' issues, and contributes to developing programs of study, and improves students' lab reports.

THE USE OF RUBRICS FOR DRAWING GRAPHS

By referring to examples from the literature on the use of rubrics for drawing graphs and to the advantages, as well as disadvantages, of using rubrics, this part of the study looks into how rubrics could be used in teaching students how to draw graphs and in assessing these drawings.

The Use of Rubrics for the Assessment of Students' Graph Drawings

In many subjects of physics classes, graphs are used in showing the relationship between variables. Helping to visualize verbal, numerical, and algebraic expressions, graphs make it easier to understand the relationship between difficult concepts (Çelik & Sağlam-Arslan, 2012). Effectively summarizing complex information or relations, graphs are used frequently in social sciences such as statistics, economy, sociology, and political science to show and interpret relationships (Bayazıt, 2011; Ozgun-Koca, 2001). Students see graphs not only in their educational life, but also in relation to other issues (Aberg-Bengtsson & Ottosson, 2006). To be able to read graphs, one needs general reasoning skills in addition to the scientific content knowledge on the information given in those graphs (Wang et al., 2012). Students with higher logical thinking skills were seen to be more successful at interpreting graphs compared to students

with lower logical thinking skills (Bektaslı & White, 2012). Studies on graph reading, interpretation, and drawing show that students from primary school through university, and in different disciplines, face difficulties and have issues with graphs (Coştu et al., 2017; Ercan et al., 2018; Ergül, 2018; Hattikudur et al., 2012; İnanç, 2019; Lowrie & Diezman, 2007; Parmar & Signer, 2005; Seçken & Çelik, 2021; Tairab & Khalaf-AlNaqbi, 2004; Yabanlı, Yıldırım & Günaydın, 2013). Graphs, which are an indispensable part of experiments, are also indispensable to lab reports (Hernandez & Korzun, 2021; Mckenzie & Padilla, 1986). In the literature, it was shown that students have difficulties in reading, interpreting, and drawing graphs on different topics such as kinematics (Bektaslı & White, 2012; McDermott et al., 1987; Beichner, 1994; Phage et al., 2017; Eryılmaz-Toksoy, 2020), standing waves and rational motion (Nixon et al., 2016), heat and temperature (Aydın, 2018), and energy graphs on mass-spring systems (Karal-Eyüboğlu, 2020). Kinematics include more graphs compared to other topics. Therefore, there are more studies on graphs in kinematics (Planinic et al., 2012).

Graphs should be used if one wants to determine the tendencies and relationships between numerous and complex data. Graphs could be line, pie, or bar graphs depending on the kind of data to be presented (Slutsky, 2014). Graph lines are frequently used in physics to show the relationship between different variables and to ensure concepts are understood more clearly. In order for line graphs to be used as a common language to show the relationship between two continuous variables, students should have certain abilities and use them appropriately (Gültekin, 2009). Studies on graphs show that students have more difficulty with line graphs compared to other graph types (Namal, 2022). Studies on the relationship between drawing graphs and graph reading-interpretation skills show that students need to have various skills and knowledge in order to draw a line graph fully correctly. For drawing graphs in physics, one uses not only field knowledge of physics, but also mathematical knowledge related to the characteristics of graphs. Moreover, drawing a line graph contains numerous steps such as determining and tagging the appropriate axis for the variables, scaling the axis, data entry, forming a point for the data, and drawing the graph curve by connecting these points (Gültekin, 2009; McKenzie & Padilla, 1986). At the end of this process of drawing a graph, which includes various skills and knowledge, it would not be appropriate to assess students' graphs simply as correct or incorrect. Studies on graphs show that rubrics are used in interpreting graphs (Boote, 2000), assessing the drawn graphs (Bahtaji, 2020; Know, 2002; Vučeljić & Šuškačević 2016; Yayla & Özsevgeç, 2014; Yeltekin Atar & Aykutlu, 2023) and teaching graphs (Angra & Gardner, 2018). A graph rubric designed to assess graphs enables them to be assessed quickly, systematically, and in a fit-for purpose way. Moreover, a graph rubric can transmit expectations about a well-designed graph for students in addition to providing communication for learning outcomes (Angra & Gardner, 2018). In their study, Angra & Gardner (2018) designed a rubric to use to teach, learn, and research about graphs. This rubric could be used to facilitate teaching and assessing graphs that summarize data, to give formative and summarizing feedback on graphs to students, and to assess students' experimental skills in relation to graphs. Boote (2000) examined what affected students' interpretation of graphs and how their interpretation was affected by the levels of graph questions. In this study, data – students' answers which they gave vocally for the questions – were gathered by using a graph drawing scoring rubric. At the end of the study, it was determined that students' answers changed according to the level of the graph question. In the literature, rubrics were used in assessing students' graph drawings in different disciplines.

Graphs are frequently used, especially in making kinematics more comprehensible (Bektaşlı & White, 2012; Demirci & Uyanık, 2009). Studies using rubrics in assessing students' graph drawings are mostly on graphs about kinematics. Yeltekin Atar and Aykutlu (2023) used an assessment rubric for the scor-

ing of high school students' graphs on force and motion. The rubric used in this study to score students' graph drawings was designed based on Tarakçı (2016). Additional categories and criteria were added to Tarakçı's evaluation criteria, and expert opinion was sought for this newly designed rubric. Scores obtained through this rubric was used in comparing students' scores in their graph reading-interpretation. The rubric used in this study is given in Table 1.

Table 1. Rubric for evaluating drawings of graphs

Evaluation Criteria	Categories	Score
Naming the Axes According to Variables	Correct (C): Naming both axes correctly and writing down the units of physical qualities representing the axes in parentheses.	2
	Partially Correct (PC): Naming only one of the axes correctly or not writing down/partially writing down/incorrectly writing down the units of physical qualities of the axes.	1
	Incorrect (I): Naming both axes incorrectly or failing to name either axis.	0
	Blank (B): Leaving the question blank, lack of any drawn graph.	0
Writing Down the Data on the Axes of the Graph	Correct (C): Correctly writing data on both axes.	2
	Partially Correct (PC): Writing down only one data group correctly and writing down the other incorrectly.	1
	Incorrect (I): Incorrectly writing down data on both axes or failing to write down any data.	0
	Blank (B): Leaving the question blank, lack of any drawn graph.	0
Creating a Point on the Axes of the Graph	Correct (C): Correctly intersecting the data on the ''y'' axis with data on the "x" axis and creating a point.	2
	Partially Correct (PC): Correctly intersecting the data of only one of the axes and making a mistake in the other one.	1
	Incorrect (I): Incorrectly intersecting data in both axes.	0
	Blank (B): Leaving the question blank, lack of any drawn graph.	0
Drawing the Curve of the Graph	Correct (C): Drawing the whole curve of the graph appropriately for the question.	2
	Almost Correct (AC): Appropriately drawing at least 3-time intervals of the 4-time-interval part of the curve of the graph or at least 2 time intervals of 3-time-interval part.	1.5
	Partially Correct (PC): Appropriately drawing at least 2-time intervals of the 4-time-interval part of the curve of the graph, or at least 1 time- interval of the 3-time-interval part of the curve of the graph.	1
	Incorrect (I): The whole curve of the graph being inappropriate.	0
	Blank (B): Leaving the question blank, lack of any drawn graph.	0
Maximum Score to Get for Each Question		8

The rubric used in this study has four sub-categories, namely, naming the axes according to variables, putting down data on the graph axes, forming a point on the graph axes, and drawing the graph curve. Each category was also further divided into correct, partially correct, incorrect, or blank. A rubric designed for the velocity-time, acceleration-time, and position-time graphs in kinematics would be used in the assessment of graph drawings, in ensuring an objective and consistent scoring, and in examining in detail students' problems when it comes to drawing graphs. In the study of Yeltekin Atar and Aykutlu

(2023) study, the "drawing the graph curve" criterion was categorized differently to other criteria, and "almost correct" was added as an additional category. When the graphic drawings of the students were examined, it was seen that they can include more than one type of motion such as constant speed, uniform acceleration, and uniform deceleration. If students had been asked to draw a graph showing uniform velocity motion for one type of motion, then the "drawing the graph curve" criterion could have been scored and scaled differently, because the drawing of a graph line differs according to different time intervals. Since the drawing of a graph varies according to motion types, this evaluation criterion was scaled differently to others. All of these show that the rubric has been designed in detail and according to the graph drawings students were asked to draw. Concerning the motion on an object, at the secondary education level, the content of force and motion includes information only on the characteristics of constant velocity motion and on graphs on constant velocity motion (MEB, 2018). Secondary school students are not taught some of the content covered in some of the evaluation criteria in the rubric. In this regard, it is believed that the rubric developed by the Yeltekin Atar and Aykutlu (2023) is not appropriate to use when evaluating secondary education students' graph drawings on constant velocity motion. Considering the performance criteria in the rubric were defined based on the taught and evaluated performance (Tierney & Simon, 2004), it can be argued that the rubric used in the study may not be appropriate at all class levels (Sadler, 2009a). If one wants to use this rubric to assess secondary school students' graph drawings of constant velocity motion, then the evaluation criterion for the drawing of the graph curve should be revised and re-appropriated for the students' level.

When the rubric used in the study was examined, it was seen that the "writing down data on the graph axes" was scored only according to the accuracy of the student's answer. Questions about velocity-time, acceleration-time, and position-time graphs, which are included in kinematics, can be solved by using the area and slope information, as well as the formulas/equations within physics. How students calculate the area and slope in kinematic graphs is also important in terms of understanding kinematic graphs (Bektaslı & White, 2012). However, in the rubric presented in Table 1, there is no scoring criterion for such a detail. How students have obtained data on the variables cannot be determined by using this scoring rubric. While it could be thought of as a deficiency, it is actually about the fact that the rubric has been designed fit for purpose. In line with the purpose of the study, it is seen that the graph drawings of the students were scored using the rubric developed according to the expected graph features. In the study, it was seen that what kind of information the students used while making their graphic drawings was not examined. Therefore, there is no evaluation criterion in the rubric which assesses which information the students have utilized to draw their graphs. It can be said that when evaluating students' graph drawings, the rubric used in the study cannot distinguish between correct graph drawings which have used different strategies. A similar problem with the use of rubric can be seen in Hull et al.'s study (2013), in which it was seen that rubrics could not differentiate between two students' solutions on the same kinematic problem in terms of their expertise. In the study, the rubric was used to ensure students' graph drawings were assessed in detail according to the time intervals of the graph in an objective and consistent manner. There was no category concerning the explanation about the title of the graph in the rubric used in the study. While it may not be necessary to have such an explanation in this study, it would be necessary for other studies to have an explanation that asks students for the title of the graph. The category concerning the explanation could be scaled as "those who explained correctly," "those who explained partially," and "those who explained incorrectly;" also different scorings could be used. It is believed that this explanation would help the graph to be understood fully, because when dealing with graphs explaining the relationship between data obtained from experiments in lab studies, such

explanations are important to indicate the accuracy of the drawing and why the graph has been drawn in the first place.

Another rubric for assessing students' graph drawings was designed by Bahtaji (2020). In his study, Bahtaji (2020) tried to determine the effects of supportive graph interventions on students' graph drawing skills and conceptual understandings. The graph drawing skill test used in the study was designed to measure students' graph drawings in relation to Newtonian mechanics. Whether these drawings were "high level graphs" or "low level graphs" was classified by the use of this rubric.

Graphs are formed to show the relationship between concepts in laboratories, which are an important component of physics classes. Nixon et al. (2016) examined undergraduate students' graphs and their interpretations during two different physics lab activities on standing waves and the rotating motion. Graphs drawn by students were assessed in three different ways in the study. One of these assessments was done by employing the graph quality rubric. The graphs students drew after two lab activities on standing waves and the rotating motion were assessed by using a graph quality rubric. The graph quality rubric was designed by taking into account the data set students used to draw the graph and the aims of the graph. There are categories and sub-categories on the graph quality rubric, namely, axes labeled, scale, accurate data points, and cleanliness. One of the results obtained in the study was that the most difficult aspect of forming a graph was forming the scale of the graph. In these three studies where rubrics were used to assess students' graph drawings, the rubrics contained different performance criteria since they were designed in accordance with the aim of the respective study.

The Use of Rubrics in Teaching How to Draw Graphs

Frequently used in physics classes and physics labs, graphs ensure that concepts are better understood and present brief sets of information instead of numerous data (Demirci & Uyanık, 2009). Students come across graphs in various settings in physics classes: there are graphs showing the relationship between concepts in a given topic, graphs included in the problem case given in a graphic presentation, graphs required of students for a problem case, or graphs about data obtained after a lab application. Considering the characteristics of graphs in physics classes, students not only need certain knowledge and skills, but they should also know the steps of forming a graph. Drawing a graph is considered a complex process, and in the literature, it was indicated that students had difficulties selecting the scale and the starting value, and that they had incorrect or incomplete drawings (Hernandez & Korzun, 2021; Nixon et al., 2016). There have been no studies in the literature on the use of rubrics in teaching how to draw a graph in physics education. Students may be given a rubric which contains a teachers' graph drawing indicating how to draw a graph so that the subject is summarized properly, and students learn more permanently. Students may be given a rubric after they complete their drawings, and they could be asked to assess their own drawings. As such, they could see their own mistakes after this self-assessment activity. Self-assessment or peer assessment of graph drawings would help students improve their drawings.

THE ADVANTAGES AND DISADVANTAGES OF USING RUBRICS FOR DRAWING GRAPHS

Based on the information presented in studies on rubrics, it is evident that the use of rubrics in teaching how to draw a graph or assessing graph drawings may have some advantages, as well as disadvantages. A

rubric prepared for the graphs which are used to show the relationship between the variables according to each time interval will provide a great advantage to the field educators in terms of being detailed, objective, and consistent in the evaluation of the student graphs. Rubrics that provide an objective assessment will help teachers to get feedback in determining where students have deficiencies in drawing graphs (Aslanoğlu & Kutlu, 2003). For instance, a student may enter the correct data on the graph but may fail to correctly draw the graph curve. Using a rubric may identify this more clearly and make it possible to examine the root of the problem. Rubrics also shorten the assessment time (Aydın-Gürler & Baykara, 2020). Co-designing the rubric that will be used in the assessment of the graph drawings together with the students would be beneficial for them since they would know the points that they need to pay attention to while drawing the graph. In the literature, it was noted that rubrics co-designed with students helped decrease students' anxiety and increase their success (Andrea & Du, 2005; Kocakülah, 2020). Giving students a rubric on graphs before their drawing would help them understand what is expected of them, because some graph components may seem insignificant for students, and they may fail to indicate them on the graph. For instance, they may fail to indicate the variables or the units of these variables. Such a failure would prohibit a complete assessment of the student's graph drawing. Some studies argued that giving rubrics to students beforehand would hinder students' creativity (Güneş & Soran, 2013). Providing the students with a graph before they draw the graph may be thought of as a disadvantage that would hinder their creativity. However, drawing graphs that show the relationship between variables in physics does not necessarily require creativity. Students may employ different calculation methods in finding data for variables in their graph drawings. For example, data on the variables in drawing kinematic graphs can be found by using equations; but they can also be determined by calculating the area beneath the graph curve or calculating the slope of the graph curve. It is thought that including these different strategies in the rubrics in students' graph drawings may be beneficial for the students to see or remember different solutions before drawing the graph; but it is also thought that this may negatively affect the students' development of their own strategies while drawing graphs, because it may prevent them from reasoning and making connections between subjects. Moreover, it is thought that it is not easy to design a rubric that would assess different strategies while assessing graph drawings because it is time consuming and requires expertise. Students may be asked to self-assess or assess their peers so that they can see their mistakes and points to improve (Wolf & Steven, 2007). There are studies showing that self-assessment activities which contain rubrics increase students' performance (Etkina et al., 2006; Faletič & Planinšič, 2020). If self or peer review will be utilized as a means of improving graph drawings, students' performance levels should be taken into account. If rubrics were to be used to facilitate students' learning, students should understand the difference between their own performance levels and the performance level expected of them (Lipnevich, et al., 2014). Low-performing students may not be able to understand an expectation regarding a high performance (Sadler, 1983). If peer-review is considered when using the rubric, then one should also consider performance levels as well as the relationship between students, because students with low self-esteem or students who are unfamiliar with peer-review may not perform as expected in such an activity. However, if students are supported and if their reservations are addressed properly, their learning may improve (Sadler, 2009b). Clearly laying out a teacher's expectation, rubrics provide feedback for students (Andrade, 2000). However, rubrics cannot provide feedback on content. Therefore, teachers' guidance is significantly important (Etkina et al., 2006). When students are aware of the expectations, they can be motivated to meet them. Using rubrics when teaching how to draw a graph could give students formative feedback about their drawings (Andrade, 2005).

CONCLUSION

Graphs are used to show the relationship between concepts in various topics and lab applications in physics classes. Based on the findings of this study on rubrics, it can be argued that rubrics may be used by physics teachers/educators in assessing students' graph drawings and in teaching them how to draw graphs. Rubrics, which allow students' performance to be evaluated according to certain criteria, can be used in the assessment of students' drawings, and the points where students make the most mistakes can be identified. By using rubrics in the assessment of students' graph drawings, these drawings are assessed in a fair and consistent manner and the assessment is completed in a short time. Co-designing the rubric to be used in the assessment of the graph drawings together with the students will be beneficial for the students because they can see what is expected of them thereby reducing their anxiety. In addition to the assessment of graph drawings, rubrics can also benefit teachers when used in teaching graphs. Rubrics can be developed in a collaborative environment in order for students to better understand the graphs in a given subject and to realize their deficiencies in drawing graphs. In addition, having self-assessment and peer assessment in assessing graph drawings would provide feedback for students so that they can see their deficiencies in drawing graphs, which, in return, will support the improvement of their drawings. A teacher who wants to employ rubrics in the evaluation of students' graph drawings for physics subjects should check the content of the subject with the graph, the graph characteristics and whether the rubric prepared for the graph contains evaluation criteria for the specified characteristics. In addition, if rubrics for the teaching of graph drawing are to be used, care should be taken to ensure that the explanations regarding the performance criteria are appropriate for the level of the students and are comprehendible.

REFERENCES

Aberg-Bengtsson, T., & Ottosson, T. (2006). What lies behind graphicacy? Relating students' results on a test of graphically represented quantitative information to formal academic achievement. *Journal of Research in Science Teaching*, *43*(1), 43–62. doi:10.1002/tea.20087

Aktaş, M., & Alıcı, D. (2018). Analytical rubric development for story writing: Validity and reliability study. *Mersin University Journal of the Faculty of Education*, *14*(2), 597–610. doi:10.17860/mersinefd.424198

Andrade, H. G. (1997). Understanding rubrics. *Educational Leadership*, *54*(4), 14–17.

Andrade, H. G. (2000). Using rubrics to promote thinking and learning. *Educational Leadership*, *57*(5), 13–18.

Andrade, H. G. (2005). Teaching with rubrics: The good, the bad and the ugly. *College Teaching*, *53*(1), 27–31. doi:10.3200/CTCH.53.1.27-31

Andrade, H. G., & Du, Y. (2005). Student perspectives on rubric-referenced assessment. *Practical Assessment, Research & Evaluation*, *10*(3), 1–11.

Angra, A., & Gardner, S. M. (2018). The graph rubric: Development of a teaching, learning, and research tool. *CBE Life Sciences Education*, *17*(4), 1–18. doi:10.1187/cbe.18-01-0007 PMID:30496033

Arpaguş, E. K., Ünsal, Y., & Moğol, S. (2011). The Effects of visual literacy on the success of secondary school students in spherical mirrors and lenses. [NWSA]. *E-Journal of New World Sciences Academy*, *6*(3), 1972–1981.

Arter, J. (2000). Rubrics, scoring guides, and performance criteria. Classroom tools for Assessing and improving student learning. The American Educational Research Association, New Orleans, LA, April, 1-22.

Aslanoğlu, A. E., & Kutlu, Ö. (2003). Öğretimde sunu becerilerinin değerlendirilmesinde dereceli puanlama anahtarı (rubrıc) kullanılmasına ilişkin bir araştırma. *Ankara University*. *Journal of Educational Sciences*, *36*(1), 25–36.

Aydın, A., & Tarakçı, F. (2018). The Investigation of the pre-service science teachers' abilities to read, interpret and draw graphs. *Elementary Education Online, 17*(1), 469-488. https://doi.org/. 2018.413806 doi:10.17051/ilkonline

Aydın, N. (2018). Determining of science pre-service teachers' graphic understanding and interpretation levels in context of heat and temperature. *Fen Bilimleri Öğretimi Dergisi*, *6*(1), 20–36.

Aydın Gürler, S., & Baykara, O. (2020). The use of scoring rubric in science course and student opinions on the scoring rubric: A mixed method study. *Turkish Studies -. Education in Science*, *15*(2), 673–690.

Bahtaji, M. A. A. (2020). Improving students graphing skills and conceptual understanding using explicit graphical physics instructions. *Cypriot Journal of Educational Science.*, *15*(4), 843–853. doi:10.18844/cjes.v15i4.5063

Bayazıt, İ. (2011). Prospective teachers' understanding of graphs. *Gaziantep University Journal of Social Sciences*, *10*(4), 1325–1346.

Beichner, R. (1994). Testing student interpretation of kinematics graphs. *American Journal of Physics*, *62*(8), 750–762. doi:10.1119/1.17449

Bektaslı, B., & White, A. L. (2012). The relationships between logical thinking, gender, and kinematics graph interpretation skills. *Eurasian Journal of Educational Research*, *48*, 1–20.

Boote, S. K. (2014). Assessing and understanding line graph interpretations using a scoring rubric of organized cited factors. *Journal of Science Teacher Education*, *25*(3), 333–354. doi:10.100710972-012-9318-8

Çelik, D., & Sağlam Arslan, A. (2012). The analysis of teacher candidates' translating skills in multiple representations. *Elementary Education Online*, *11*(1), 239–250.

Chasteen, S. V., Pepper, R. E., Caballero, M. D., Pollock, S. J., & Perkins, K. K. (2012). Colorado upper-division electrostatics diagnostic: A conceptual assessment for the junior level. *Physical Review Special Topics. Physics Education Research*, *8*(2), 020180. doi:10.1103/PhysRevSTPER.8.020108

Chasteen, S. V., & Scherr, R. E. (2020). Developing the physics teacher education program analysis rubric: Measuring features of thriving programs. *Physical Review. Physics Education Research*, *16*(1), 010115. doi:10.1103/PhysRevPhysEducRes.16.010115

Coştu, F., Beler, Ş., & Coştu, B. (2017). Students' graphic skills and their difficulties about photosynthesis topic. *Scientific Educational Studies,* 41-63.

Demirci, N., & Uyanık, F. (2009). The correlation between tenth grade students' understanding and interpreting graphs and their kinematics achievment. *Necatibey Faculty of Education. Electronic Journal of Science and Mathematics Education*, *3*(2), 22–51.

Docktor, J. L., Dornfeld, J., Frodermann, E., Heller, K., Hsu, L., Jackson, K. A., Mason, A., Ryan, Q. X., & Yang, J. (2016). Assessing student written problem solutions: A problem-solving rubric with application to introductory physics. *Physical Review. Physics Education Research*, *12*(1), 010130. doi:10.1103/PhysRevPhysEducRes.12.010130

Ercan, O., Coştu, F., & Coştu, B. (2018). Identifying pre-service science teachers' difficulties about graph drawings. *Kastamonu Education Journal*, *26*(6), 1929–1938.

Ergül, N. R. (2018). Pre-service science teachers' construction and interpretation of graphs. *Universal Journal of Educational Research*, *6*(1), 139–144. doi:10.13189/ujer.2018.060113

Eryılmaz Toksoy, S. (2020). Investigation of 11th grade students' knowledge about explanation of motion types and drawing, interpreting related graphs. *Bolu Abant İzzet Baysal University Journal of Faculty of Education*, *20*(3), 1423–1441.

Etkina, E., Heuvelen, A. V., White-Brahmia, S., Brookes, D. T., Gentile, M., Murthy, S., Rosengrant, D., & Warren, A. (2006). Scientific abilities and their assessment. *Physical Review Special Topics. Physics Education Research*, *2*(2), 020103. doi:10.1103/PhysRevSTPER.2.020103

Faletič, S., & Planinšič, G. (2020). How the introduction of self-assessment rubrics helped students and teachers in a project laboratory course. *Physical Review. Physics Education Research*, *16*(2), 020136. doi:10.1103/PhysRevPhysEducRes.16.020136

Gültekin, C. (2009). *Examining 9th grade students' abilities on drawing reading and interpreting of graphs about solutions and their properties.* [Unpublished master's thesis, Balikesir University, Türkiye].

Güneş, P., & Kılıç, D. (2016). Self-, peer- and teacher-assessment through rubrics. *Mehmet Akif Ersoy University Journal of Education Faculty*, *39*, 58–69.

Güneş, P., & Soran, H. (2013). Secondary school students' opinions on rubrics. *Kastamonu Education Journal*, *21*(4), 1327–1344.

Hafner, J. C., & Hafner, P. M. (2003). Quantitative analysis of the rubric as an assessment tool: An emprical study of student peer-group rating. *International Journal of Science Education*, *25*(12), 1509–1528. doi:10.1080/0950069022000038268

Hall, E. K., & Salmon, S. J. (2003). Chocolate chip cookies and rubrics helping students understand rubrics in inclusive settings. *Teaching Exceptional Children*, *35*(4), 8–11. doi:10.1177/004005990303500401

Hattikudur, S., Prather, R. W., Asquith, P. S., Alibali, M. W., Knuth, E. J., & Nathan, M. (2012). Constructing graphical representations: Middle schoolers' intuitions and developing knowledge about slope and y-intercept. *School Science and Mathematics*, *112*(3), 230–240. doi:10.1111/j.1949-8594.2012.00138.x

Hernandez, J., & Korzun, B. (2021). Methods to select the scales and starting values for axes in linear graphs. *The Physics Teacher*, *59*(6), 482–483. doi:10.1119/10.0006136

Hull, M. M., Kuo, E., Gupta, A., & Elby, A. (2013). Problem-solving rubrics revisited: Attending to the blending of informal conceptual and formal mathematical reasoning. *Physical Review Special Topics-Review Physics*. *Education Research*, *9*, 010105.

İnanç, H. (2019). Graphic Drawing skills of science teaching candidates. *International Scientific and Vocational Journal*, *3*(1), 22–30.

Jonsson, A., & Svingby, G. (2007). The use of scoring rubrics: Reliability, validity and educational consequences. *Educational Research Review*, *2*(2), 130–144. doi:10.1016/j.edurev.2007.05.002

Kan, A. (2009). Ödev ve Projeler. Eğitimde Ölçme ve Değerlendirme, Hakan Atılgan (Ed.), 4. Baskı Anı Yayıncılık, Ankara.

Karaçam, Z. (2013). Systematic review methodology: A guide for preparation of systematic review. *E-Journal of Dokuz Eylul University Nursing Faculty*, *6*(1), 26–33.

Karal Eyüboğlu, I. S. (2020). Interpretation of an energy graph for a mass-spring system by prospective science and mathematics teachers: A comparison. *Online Science Education Journal*, *5*(2), 52–59.

Kocakülah, A. (2022). Development and use of a rubric to assess undergraduates' problem solutions in Physics. *Participatory Educational Research*, *9*(3), 362–382. doi:10.17275/per.22.71.9.3

Kocakülah, M. S. (2010). Development and application of a rubric for evaluating students' performance on Newton's laws of motion. *Journal of Science Education and Technology*, *19*(2), 46–164. doi:10.100710956-009-9188-9

Kocakülah, M. S., & Aytaç, N. N. (2007). Development and use of a rubric for the assessment of students' performance in solving problems on Newton's law of motion. *AIP Conference Proceedings*, *899*, 838. doi:10.1063/1.2733579

Kwon, O. N. (2002). The effect of calculator based ranger activities on students' graphing ability. *School Science and Mathematics*, *102*(2), 57–67. doi:10.1111/j.1949-8594.2002.tb17895.x

Liew, S. S., Lim, H. L., Saleh, S., & Ong, S. L. (2019). Development of scoring rubrics to assess physics practical skills. *Eurasia Journal of Mathematics, Science and Technology Education*, *15*(4), 1–14. doi:10.29333/ejmste/103074

Lipnevich, A. A., McCallen, L. N., Miles, K. P., & Smith, J. K. (2014). Mind the gap! Students' use of exemplars and detailed rubrics as formative assessment. *Instructional Science*, *42*(4), 539–559. doi:10.100711251-013-9299-9

Lowrie, T., & Diezmann, C. M. (2007). Middle school students interpreting graphical tasks: difficulties within a graphical language. In *4th East Asia Regional Conference on Mathematics Education*, (pp. 611-617).

McDermott, L. C., Rosenquist, M. L., & Van Zee, E. H. (1987). Student difficulties in connecting graphs and physics: Examples from kinematics. *American Journal of Physics*, *55*(6), 503–513. doi:10.1119/1.15104

Mckenzie, D. L., & Padilla, M. J. (1986). The construction and validation of the test of graphing in science (TOGS). *Journal of Research in Science Teaching*, *23*(7), 571–579. doi:10.1002/tea.3660230702

Mertler, C. A. (2001). Designing scoring rubrics for your classroom. *Practical Assessment, Research & Evaluation*, *7*(25), 1–8.

Ministry of National Education (MEB). (2018). *Fen bilimleri dersi öğretim programı* [Science course teaching program curriculum]. Talim ve Terbiye Kurulu Başkanlığı.

Moskal, B. M. (2000). Scoring rubrics: What, when and how? *Practical Assessment, Research & Evaluation*, *7*(3), 1–5.

Munezero, V., Yadav, L., & Bugingo, J. B. (2022). Representation of nature of science in physics textbooks of cycle 4 fundamental schools in Burundi. *European Journal of Educational Research*, *11*(4), 2487–2496. doi:10.12973/eu-jer.11.4.2487

Namal, R. (2022). Multiple examination of graduate theses within the scope of graph literacy in social studies teaching. *Karaelmas Journal of Educational Sciences*, *10*, 87–102.

Nixon, R. S., Godfrey, T. J., Mayhew, N. T., & Wiegert, C. C. (2016). Undergraduate student construction and interpretation of graphs in physics lab activities. *Physical Review. Physics Education Research*, *12*(1), 010104. doi:10.1103/PhysRevPhysEducRes.12.010104

Oaklef, M. (2009). Using rubrics to assess information literacy: An examination of methodology and interrater reliability. *Journal of the American Society for Information Science and Technology*, *60*(5), 969–983. doi:10.1002/asi.21030

Ozgun-Koca, S. A. (2001). The graphing skills of students in mathematics and science education. *ERIC Digest.* https://files.eric.ed.gov/fulltext/ED464804.pdf

Papakonstantinou, M., & Skoumios, M. (2021). Science and engineering practices in the content of Greek middle school physics textbooks about forces and motion. *Journal of Technology and Science Education*, *11*(2), 457–473. doi:10.3926/jotse.1286

Parmar, R. S., & Signer, B. R. (2005). Sources of error in constructing and interpreting graphs a study of fourth-and fifth-grade students with LD. *Journal of Learning Disabilities*, *38*(3), 250–261. doi:10.1177/00222194050380030601 PMID:15940962

Phage, I. B., Lemmer, M., & Hitge, M. (2017). Probing factors influencing students' graph comprehension regarding four operations in kinematics graphs. *African Journal of Research in Mathematics. Science and Technology Education*, *21*(2), 200–210.

Planinic, M., Milin-Sipus, Z., Katic, H., Susac, A., & Ivanjek, L. (2012). Comparison of student understanding of line graph slope in physics and mathematics. *International Journal of Science and Mathematics Education*, *10*(6), 1393–1414. doi:10.100710763-012-9344-1

Popham, J. W. (1997). What's wrong and what's right with rubric. *Educational Leadership*, *55*(2), 72–75.

Popham, J. W. (2007). *Classroom assessment: What teachers need to know* (5th ed.). Pearson Education.

Sadler, D. R. (1983). Evaluation and the improvement of academic learning. *The Journal of Higher Education*, *54*(1), 60–79. doi:10.2307/1981645

Sadler, D. R. (2009a). Transforming holistic assessment and grading into a vehicle for complex learning. *Assessment, Learning and Judgement in Higher Education*. In G. Joughin (Ed.), *Springer Science+Business Media B.V.*, doi:10.1007/978 1 4020 8905 3_4

Sadler, D. R. (2009b). Indeterminacy in the use of preset criteria for assessment and grading in higher education. *Assessment & Evaluation in Higher Education*, *34*(2), 159–179. doi:10.1080/02602930801956059

Salazar-Torres, J., Leal, O. R., & Ortega, M. V. (2021). The rubric as an assessment tool for solving problem situations in the physics and mathematics teaching context. *Journal of Physics: Conference Series: V International Meeting of Mathematical Education*. IOP Science. 10.1088/1742-6596/1981/1/012018

Seçken, N., & Çelik, Ç. (2021). Investigating high school students' graphic interpretation skills on the subject of chemical equilibrium. *Journal of Research in Education and Society*, *8*(1), 179–204.

Slutsky, D. J. (2014). The effective use of graphs. *Journal of Wrist Surgery, 03*(02), 067-068. doi:10.1055/s-0034-1375704

Tairab, H. H., & Khalaf-AlNaqbi, A. K. (2004). How do secondary school science students interpret and construct scientific graphs? *Journal of Biological Education*, *38*(3), 127–132. doi:10.1080/00219 266.2004.9655920

Tarakçı, F. (2016). *Examining science teacher candidates' abilities on reading, interpreting and preparing of graphs.* [Unpublished master's thesis, Kastamonu University, Türkiye].

Taşar, M. F., İngeç, Ş. K., & Güneş, P. Ü. (2002). *Determining graphic drawing and understanding skills.* V. National Science and Mathematics Education Congress, ODTÜ, Ankara.

Teodorescu, R. E., Bennhold, C., Feldman, G., & Medsker, L. (2014). Curricular reforms that improve students' attitudes and problem-solving performance. *European Journal of Physics Education*, *5*(1), 15–44. doi:10.20308/ejpe.91287

Tierney, R., & Simon, M. (2004). What's still wrong with rubrics: Focusing on the consistency of performance criteria across scale levels. *Practical Assessment, Research & Evaluation*, *9*(2), 1–7.

Vučeljić, M., & Šuškačević, M. (2016). Achievements of Montenegrian high-school students in TUGK test (test of understanding graphs – kinematics). *AIP Conference Proceedings*, *1722*, 310007–3. doi:10.1063/1.4944317

Wang, Z. H., Wei, S., Ding, W., Chen, X., Wang, X., & Hu, K. (2012). Students cognitive reasoning of graphs: Characteristics and progression. *International Journal of Science Education*, *34*(13), 2015–2041. doi:10.1080/09500693.2012.709333

Wolf, K., & Stevens, E. (2007). The role of rubrics in advancing and assessing student learning. *The Journal of Effective Teaching*, *7*(1), 3–14.

Yabanlı, H., Yıldırım, B., & Günaydın, Ö. (2013). Translating from map to line graph. [ATED]. *Araştırma Temelli Etkinlik Dergisi*, *3*(1), 12–19.

Yayla, G., & Özsevgeç, T. (2014). The examination of secondary school students' graphic skills: Construction and interpretation of line graphs. K. Ü. *Kastamonu Education Journal*, *23*(3), 1381–1400.

Yeltekin Atar, B. Ş., & Aykutlu, I. (2023). High school students' user skills concerning force and motion graphs. *Gazi University Journal of Gazi Education Faculty*, *43*(1), 211–242. doi:10.17152/gefad.1205369

Yılmaz, K. (2021). Systematic review, meta evaluation, and bibliometric analysis in social sciences and educational sciences. *MANAS Journal of Social Studies*, *10*(2), 1457–1490.

ADDITIONAL READING

Al-Salmani, F., & Thacker, B. (2021). Rubric for assessing thinking skills in free-response exam problems. *Physical Review. Physics Education Research*, *17*(1), 010135. doi:10.1103/PhysRevPhysEducRes.17.010135

Anderson, T., & Svingby, G. (2007). The use of scoring rubrics: Reliability, validity and educational consequences. *Educational Research Review*, *2*(2), 130–144. doi:10.1016/j.edurev.2007.05.002

Arter, J., & McTighe, J. (2001). *Scoring rubrics in the classroom: Using performance criteria for assessing and improving student performance*. Corwin Press Inc.

Panadero, E., & Jonsson, A. (2013). The use of scoring rubrics for formative assessment purposes revisited: A review. *Educational Research Review*, *9*, 129–144. doi:10.1016/j.edurev.2013.01.002

Safadi, R. (2022). Supporting student learning from diagnosing erroneous examples when contrasting them with worked examples in the physics classroom. *International Journal of Science Education*, *44*(2), 245–270. doi:10.1080/09500693.2021.2023834

Safadi, R., & Saadi, S. (2021). Learning from self-diagnosis activities when contrasting students' own solutions with worked examples: The case of 10th graders studying geometric optics. *Research in Science Education*, *51*(2), 523–546. doi:10.100711165-018-9806-8

Shafer, W. D., Swanson, G., Bene, N., & Newberry, G. (2001). Effects of teacher knowledge of rubrics on student achievement in four content areas. *Applied Measurement in Education*, *14*(2), 151–170. doi:10.1207/S15324818AME1402_3

KEY TERMS AND DEFINITIONS

Drawing a graph: Visual figure formed to make the relationship between data clearer by presenting it in brief.

Graph Reading-Interpretation: Making sense of the relationships between data, explaining, and interpreting graphs.

Graph: A visual that shows the relationship between variables and presents data in brief.

Kinematics: Part of mechanics, a sub-branch of physics; kinematics tries to define an object's motion within space-time without taking into account the factors causing the motion.

Line Graph: A type of graph that shows the relationship between dependent and independent variables.

Physics Education: A branch of science which researches how to learn and teach physics, in addition to training physics teachers.

Rubric: A measurement tool used in scoring works based on certain criteria.

Chapter 3
Fostering Entrepreneurship Education by Improving Assessment Rubrics for Entrepreneurship Competence

Minna Hämäläinen

https://orcid.org/0000-0001-7677-1926
LUT University, Finland

Anu Raappana

https://orcid.org/0000-0002-6269-6934
LUT University, Finland

ABSTRACT

This chapter describes a tool for assessing the development of learners' entrepreneurship competence and its construction process. Competence in entrepreneurship is promoted at all school levels, but there are very few tools for teachers to assess the level or development of students' entrepreneurship competence. It can also be very unclear to teachers what should be assessed and when to assess a student's entrepreneurship competence. The rubric-based tool allows teachers to assess the development of a learner's entrepreneurship competence in an easy and simple way. The rubric also makes it possible to develop a teacher's own skills, as it gives an idea of what is being assessed when it comes to entrepreneurship competence. The tool was developed in a participatory process with teachers, students, and researchers. Another special feature of the presented tool is its internationality and the common contribution of different cultures to the development of the assessment tool.

INTRODUCTION

Entrepreneurship skills and knowledge are seen as a transversal competence needed in working life and teachers should promote the development of entrepreneurship related skills at all school levels (Baciga-lupo et al., 2016; European Commission, 2012; Bourgeois et al., 2016; Cotoi et al., 2011). The overall aim is to develop the ability of students to act responsibly, to be active and creative, and to be able to seize opportunities, take controlled risks, and plan and manage projects of suitable sizes (Bacigalupo et al., 2016; Henry et al. 2005; Henry and Lewis, 2018; Fiet, 2001: Komarkova et al. 2015). However, there are a lack of guidelines regarding how to evaluate the state and development of entrepreneurship competence amongst the students (e.g., Schelfhout et al. 2016).

The purpose of this chapter is to describe and analyze the assessment rubric related to the evaluation of entrepreneurship competence of students. The rubric is a part of a larger assessment project, which consists of building general assessment rubrics for selected competence and building a self- assessment tool for students about entrepreneurship competence. Practical entrepreneurial activities including assessment rubrics were also provided for teachers (Table 1). The actual assessment tool is an online tool. It does not use a central database, so therefore all the data is stored on the users' browser. However, the users can share their data between their devices by exporting and importing. The results are calculated automatically and can be downloaded by each student.

Table 1. Description of the assessment tool

ASSESSMENT TOOL	
General assessment rubrics for entrepreneurship competence.	Self- assessment tool for students about entrepreneurship competence.
Guided practical entrepreneurial activities for teachers. Assessment rubrics for entrepreneurial activities.	

The assessment rubric presented in this chapter is intended for primary and secondary school teachers. However, the same challenge in assessing the development of entrepreneurship competencies occurs at all levels of education, from early childhood education to higher education (e.g., Almeida and Buzady, 2019; Kyndt and Baert, 2015; Lans and Gulikers, 2010; Mitchelmore and Rowley, 2010). Hence, this chapter informs the debate on entrepreneurship competencies in general.

The objective of this chapter is to describe the background and theoretical framework of the assessment tool and rubric mentioned above. In addition, the participatory process of developing the rubric is described in detail. The Entrepreneurship Competence Framework (Bacigalupo et al., 2016) was used as a guideline in developing the content of the tool but the rubric was built together with the teachers and thus is based on their needs.

The actual tool is multilevel and wide-ranging and covers both the teacher and student perspective on assessing the development of entrepreneurship competence. In this chapter, we mainly concentrate analyzing the rubric from the teacher's point of view. However, the student self-assessment tool is also shortly described because it is tightly connected to the use of the rubrics. This chapter first describes the theoretical background of the tool and after that presents the tool and rubric development process. A discussion of the tool, the rubric and their usefulness conclude this chapter.

BACKGROUND

The majority of EU countries have included entrepreneurship education in their national curriculums (Bourgeois et al., 2016) and especially at the EU-level there are guiding policy guidelines and documents that encourages teachers to implement entrepreneurial aspects in their teaching (European Commission, 2013). Some of the EU countries also have national or regional curriculums or strategies which address entrepreneurship education (Bourgeois et al., 2016). In its broader meaning, entrepreneurship education aims at creating and enhancing a student's ability to act responsibly, to be active and creative and to seize opportunities, to take controlled risks, and plan and manage projects of suitable sizes (European Commission, 2013; Henry et al. 2005; Fiet, 2001, Jones and Iredale, 2010). Adopting entrepreneurship education as pedagogy also allows greater student ownership of the learning process (Jones and Iredale, 2010).

Previous research has found that students use rubrics to support their own learning and performance (Andrade and Du, 2005). From the teacher's point of view the rubrics help in describing the course expectations and grading the courses (e.g., Andrade and Du, 2005; Andrade, 2000; Goodrich, 1997; Moskal, 2003; Popham, 1997). The development of entrepreneurship skills is often assessed through a variety of tasks and reports completed by the students such as business plans or business reports (e.g. Babatunde et al., 2021; Babatunde and El-Gohary 2019; Pittaway and Edwards 2012). For example, business reports are concrete outputs and are easy to score and evaluate. However, they are not suitable for assessing the development of individual competence in entrepreneurship.

MAIN FOCUS OF THE CHAPTER

The focus of this chapter is on description of the development process of general assessment rubrics for competence in entrepreneurship. With the assessment rubric it is possible for teachers to assess the development of a students' entrepreneurship competence and at the same time develop their own understanding and abilities related to teaching entrepreneurship skills and knowledge.

Issues, Controversies, Problems

Various materials, methods and tools have been provided to assist teachers in implementing entrepreneurial aspects in the curriculum (Bourgeois et al., 2016; Johansen and Schanke, 2013; Johansen and Somby, 2015). Earlier studies have confirmed that teachers implement entrepreneurial aspects in their teaching independently, but there are differences in the number and variation of entrepreneurship education methods used by teachers (Ruskovaara, 2014). Notwithstanding the large number and variation of materials, guiding documents and provided methods, in some cases teachers may have difficulties to implement entrepreneurial methods in their teaching (e.g., Figueiredo-Nery and Figueiredo, 2008; Gustafsson-Pesonen and Remes 2012; Seikkula-Leino 2011; Sommarström et al. 2021).

However, there is no clear definition of entrepreneurship competence in the field of entrepreneurship education research (e.g. Bird 1995; Bridge et al. 2003, 37-39; Fastré and Van Gils 2007). There is some ambiguity between concepts on many different levels (Lackeus and Sävetun 2018; Mitchelmore and Rowley 2010). The concept of competence refers to the knowledge, skills, attitudes, values, and behaviours that people need to successfully perform a task (Brophy and Kiely 2002; Le Desit-Delamare and Winterton 2005; Man, et al 2002). In this chapter, we use the concept of entrepreneurship competence to describe

an individual's entrepreneurship knowledge and skills. With the plural word competences, we refer to the EntreComp framework's sub-competencies that serve as the basis for the rubric presented in this chapter.

In previous research literature the concept of entrepreneurship competence is associated with the characteristics of an entrepreneur, such as creativity, innovativeness and risk-taking (Bacigalupo et al. 2016; Bridge et al. 2003, 37-39; Gibb 2002; Schumpeter 1982, 74-91). Moreover, in the previous research literature the concept of competence has several different meanings (e.g. Fastré and Van Gils, 2007; Le Deist and Winterton 2005). There are two forms of the concept which causes some also confusion: competence and competency (e.g. Bergevoet & Van Woerkum 2006; Garavan & McGuire 2001). The terms "competence and "competency" are often used interchangeably. However, competence often refers to performance and competency refers to, for example, the professionalism of an individual. It is something that is needed for successful performance (e.g., Boyatzis 2008; Le Deist & Winterton 2005; Woodruffe 1993).

To make this situation even more confusing, there is a lack of a definition of entrepreneurship competence for situations where adolescents are the target group of examination. For youth, entrepreneurship competence can be seen as a latent quality that emerges from an individual's knowledge, skills and attitudes (e.g. Athayde 2009; Bacigalupo et al. 2016; Obschonka 2016). Adolescents are rarely the target group for research concerning entrepreneurship competence (e.g. Aamir et al. 2019; Lackeus and Sävetun 2018; 2013; de Lourdes Cárcamo-Solís et al. 2017) and for that reason, the previous literature does not provide a clear picture of what is meant by entrepreneurship competence when the target group are young learners.

SOLUTIONS AND RECOMMENDATIONS

In the case of entrepreneurship competence in particular, a rubric is a way to promote both students' and teachers' understanding of the subject being assessed, i.e. the development of young learners' entrepreneurship competence. Previous research literature has found that the rubric is beneficial to both the student and the teacher (e.g. Bresciani et al., 2004). Rubrics are used to enhance the objectivity in assessment from the teachers' point of view (e.g. Stevens and Levi 2005) and on the other hand make the goals of the study programme clear for the student. In addition, the rubric may help learners to take responsibility of the studies when they know the assessment criteria beforehand (e.g. Reddy and Andrade, 2010). Some examples can be found in the previous research literature specifically for rubrics created for the evaluation of entrepreneurship readiness. However, these relate to the higher education sector and the promotion of entrepreneurship skills under certain degree programmes (e.g. Cuenca et al., 2016; Ferreras-Garcia 2019; Reddy and Andrade, 2010; Tahahashi and Kiyousumi, 2021).

The development of the rubric and the assessment tool was an intensive participatory process (e.g. Kujala 2003) that was carried out in collaboration with teachers. During the process all the steps of the process were documented. All the steps were validated by experts, in processes carried out by teachers around Europe. The phases of the tool building process partially overlapped, although they are presented here as a progressive process. The idea of an inclusive process was to improve the user-friendliness and technical reliability of the assessment tool.

The main participants in the rubric development process were teachers (from Italy, Portugal, Poland and Turkey), experts in entrepreneurship education (from Finland and Sweden), experts in education and educational innovation (from Spain and Italy), and students. Each group had its own role in the

process. The different roles of the actors are described in this process description in the stages in which each actor was involved. In the following, the main features of the practical steps in the process and the outcome of each phase are described.

There were six main phases in the rubric (and the assessment tool) development process. These phases were:

1. Selection of the background reference (project group)
2. Selection of the sub-competences, i.e., the core of the assessment rubric (teacher together with the project group)
3. Setting the criteria for chosen three competencies and building the general indicators for assessing competencies (teachers with the project group)
4. Developing a students' self-assessment tool
5. Pilot testing in 4 European schools
6. Training course for teachers

Selection of the Background Reference

The first step in the process was to choose a frame of reference as a basis for discussing competencies and developing a rubric. The EntreComp Framework (Bagicalupo et al. 2016.) was chosen as background reference because of its comprehensiveness. The purpose of the framework was to identify the key components of entrepreneurship competence. Hence, the EntreComp framework provided a shared conceptual model for the development process.

The EntreComp framework consists of three competence areas (Ideas and opportunities; Resources and Into action). These three competence areas are divided into 15 sub-competences. The framework also includes a number of learning outcomes to suggest what all European citizens should know, understand and be able to do to demonstrate a certain level of proficiency in terms of entrepreneurship competence (Bacigalupo et al. 2016).

Selection of the Sub-Competence

Among the fifteen sub-competence of the EntreComp framework, teachers selected the ones they found most difficult to assess. The selected sub competences create the core of the assessment tool. Hence, the selection of the sub-competences is based entirely on teachers' needs. This phase was attended by teachers from four different schools. The selected sub-competences were 1) Vision (Ideas and Opportunities), 2) Self-awareness and self-efficacy (Resources) 3) Working with others (Into action). It should be noted that these three sub-competencies do not provide a comprehensive picture of entrepreneurship competence.

Vision refers to the ability to perceive the future and opportunities in one's own operating environment. Identifying and seizing opportunities is at the heart of entrepreneurship and opportunities and their detection have been one of the key themes of entrepreneurship research (e.g. Baron 2006; Short, et al. 2010.) In the EntreComp Framework the concept of vision refers to individuals' ability to *work towards a vision of the future.*

Self-awareness is the extent to which individuals sees themselves as others see them (e.g. Fletcher and Bailey, 2003; Ustav and Venesaar 2018). In the entrepreneurship education research literature, the concept of self-efficacy refers to the judgements people formulate about their capacity to act in specific

situations. Moreover, self-efficacy is a personal attribute which is found to be an important antecedent to entrepreneurship (e.g. Bandura 1977, 1997; Boyd & Vozikis, 1994; Markman & Baron, 2003; McGee et al. 2009). Self-awareness and self-efficacy have been studied extensively. There are several different meters to measure them. Previous research was used to prepare this rubric. For example, Scholz et al. (2002) have measured individual self-efficacy. The data from their research is international and therefore supports the construction of this rubric well. In the EntreComp Framework sub-competence self-awareness and self-efficacy are defined as *believing in oneself and constantly evolving.*

The third chosen sub-competence *working with others* refers in the EntreComp Framework to personal skills in building teams, working together and networking. In previous entrepreneurship research literature these attributes relate to teamworking skills and the ability to work towards common goals (e.g. Bridge et al. 2003; Silveyra et al., 2021). Teamworking skills have been seen in the previous literature as soft entrepreneurial skills. The ability to work in a team is seen as a skill that applies to all areas and all work tasks in one way or another. (e.g. van Dam et al., 20120; Robinson and Stubberud, 2014).

Setting the Criteria for Chosen Sub-Competences and Building the General Indicators

Together with the teachers the assessment criteria were set for the chosen sub-competences. Table 2. describes the assessment criteria for the selected sub-competences.

Table 2. The assessment criteria for the selected sub-competences

Sub-competence	The criteria to evaluate	Means of evaluation
Vision: imagining the future, thinking strategically, developing future scenarios.	Number of ideas for the future imagined. Easy to imagine future ideas. Novelty of ideas imagined . Argumentation of ideas. Clarity in plans/schedule developed. Documentation of ideas development. Feasibility/realism on the future scenarios imagined. Ability to consider risks and potentialities.	Brainstorming experiences. Plans for future projects (text or multimedia format). Peer-to-peer assessment/evaluation.
Self-awareness and self-efficacy: believe in oneself and keep on developing.	Awareness of everything in an individuals' environment. Consciousness of one's inner feelings. Consciousness of what is happening. Ability to reflect on one's life situation. Concern about what other people think of oneself. Confidence in having enough persistence to achieve goals. Confidence in overcoming opponents. Confidence in handling and overcoming obstacles. Confidence in solving difficult problems.	Report in text or multimedia format, Individual or in groups. Self-report of behavior during a particular event or in a day class (assessed by peers, students, or the teacher). Plans for carrying out projects (including self-efficacy report on each plan created). Self-reports about students' academic performance (assessed by peers, students, or the teacher). Existing online self-assessment questionnaires (students receive automatic feedback).
Working with others: building teams, working together and networking.	Work together and co-operate with others to develop ideas and turn them into action. Network. Solve conflicts and face up to competition positively when necessary.	Assessed either in a group situation or by analyzing each participant separately in a specific activity. Analysis done by the teacher or by a peer. Self-assessment.

Pilot Testing

The assessment tool was disseminated and piloted at online events in 4 European schools in Italy, Portugal, Poland and Turkey. The original idea for piloting with teachers and students was to do so at face-to-face events, but because of covid-19 restrictions the piloting had to be delivered online. Finally, an e-tool was piloted with approximately 80 students testing six different activities, 2 from each dimension.

In the piloting the virtual tool and assessment rubrics were piloted with 24 piloting teachers from Italy, Portugal, Poland and Turkey in a face-to-face training course. After the pilot testing, the rubrics were compiled into coherent tables, with a separate rubric for each sub-competence. These rubrics are presented in Tables 3, 4 and 5.

Table 3. Assessment rubric for the sub-competence of vision

Criteria to assess the sub-competence	Entrepreneurship Competence Assessment Framework			
	Strong competence	**Good competence**	**Requires guidance**	**Requires strong guidance**
Number of ideas for the future imagined.	The student always proposes more ideas than needed on every occasion.	The student proposes as many new ideas as needed on every occasion.	The student does not always propose as mas many ideas as needed on every occasion.	The student has insufficient imagination or has no new ideas.
Easy to imagine future ideas.	Fast and natural proposal of new ideas.	Focused, systematic and successful proposing of new ideas.	Slow but constant and successful proposing of some new ideas.	Slow and unsuccessful proposing of new ideas.
Novelty of ideas imagined.	The ideas proposed are completely new and unknown.	The ideas proposed are based on other ideas from external sources but are new proposals.	The ideas proposed are based on other ideas from external sources.	Cannot imagine any new ideas.
Argumentation of ideas.	The student argues, debates and defends ideas clearly.	The student argues and debates ideas.	The student argues ideas.	The student cannot argue or debate ideas.
Clarity in plans and schedule developed.	The student has developed a clear, detailed and realistic plan and schedule.	The student has developed a clear and realistic plan and schedule.	The student has developed a plan and schedule, but it is not completely realistic or clear.	The student has not created any plan or schedule, or the plan or schedule is not feasible.
Documentation of ideas development.	The student presents all the sources used to define the plan and knows each one of the sources very well.	The student presents all the sources used to define the plan and can comment on some of them.	The student presents some sources used to define the plan.	The student does not present any of the sources used to define the plan.
Feasibility and realism of the future scenarios imagined	The scenarios proposed are completely realistic.			The scenarios proposed are not realistic at all.
Ability to consider risks and potential.	Apart from normal risks, the student has even considered the less probable risks and potential associated with proposal.	The student has taken into account the most probable risks and potential associated with the proposal.	The student has taken into account some risks and potential associated with the proposal.	The student has not taken into account any risks or potentiality associated with the proposal.

Table 4. Assessment rubric for the sub-competence of self-assessment and self-efficacy

Criteria to assess the sub-competence	Entrepreneurship Competence Assessment Framework			
	Strong competence	**Good competence**	**Requires guidance**	**Requires strong guidance**
Awareness of everything in its environment.	The student pays attention to the details of what appears in the surrounding world. The student understands the complex and wide relationships between different elements.	The student pays attention to the surrounding world but misses some elements. The student understands the relationships between different elements.	The student pays attention to some of the details in the surrounding world but does not relate to them or understand them as a complex context.	The student pays no attention to the surroundings, shows no interest in what is around him/her.
Consciousness of what is happening.	The student describes in detail the situation in which student finds him/herself and understands the process that happen in that situation in a complex way. The students pay close attention to details.	The student describes the situation in which he/she finds him/herself and understands the processes that take place in it.	The student knows that something is happening, he/she describes it superficially, but is not able to describe it completely.	The student has problems in understanding and describing different situations.
Consciousness about the image she/he projects.	The student knows what image she/he projects of her/himself and understands that the image includes not only that it shows what he/she does, but also the relationships and people around him/her.	The student believes she/he projects an image of her/himself based on what others see of it and what it does.	The student believes she/he projects an image of her/himself based on what others see of him/her.	The student has problems in knowing what image she/he projects of her/himself.
Concern about what other people think about him/her.	The student is aware of the image she/he projects of her/himself. When receiving constructive feedback, the student can make conscious change that image.	The student is aware of the image she/he projects of her/himself. The student does not modify that image.	The student is partially aware of the image she/he projects of her/himself.	The student pays no attention to what image he/she projects to others. At the opposite extreme, everything she/he does is done by thinking about what others think of him/her.
Consciousness of student's own inner feelings.	The student knows his/her own feelings and preferences and their causes and expresses them assertively.	The student know his/her feelings and preferences and expresses them.	The student does not really show nor recognize his/her own feelings and preferences but always looks for others' feelings and/or preferences for reference.	The student does not understand or express his/her own feelings or preferences.
Confidence on having enough persistence to achieve goals.	The student is convinced that he/she can persist until achieving goals, even big or tough.	The student believes his/hers potential to persist until achieving most of the goals.	The student doubt about on his/ hers potential to persist until achieving most of the goals.	The student is convinced that he/she cannot persist until achieving the goals.
Confidence on handling and overcoming obstacles.	The student is convinced that he/she can handle and overcome obstacle, even unexpected or difficult.	The student has conviction of his/hers potential handling and overcoming majority of potential obstacles.	The student believes on his/ hers potential handling and overcoming some easy obstacles.	The student is convinced that he/she cannot handle any obstacle.
Confidence on overcoming opponents.	The student is convinced that he/she can any opponent.	The student has conviction of his/hers potential overcoming potential opponents.	The student has conviction of his/ hers potential overcoming some easy opponents.	The student is convinced that he/she cannot overcome any opponent.
Confidence on solving difficult problems.	The student is convinced that he/she can solve any problem, even unexpected or difficult.	The student has conviction of his/ hers potential solving majority of potential problems.	The student has conviction of his/ hers potential solving some easy problems.	The student is convinced that he/she cannot solve any problem.

Table 5. Assessment rubric for the sub-competence of working with others

Criteria to assess the sub-competence	Entrepreneurship Competence Assessment Framework			
	Strong competence	**Good competence**	**Requires guidance**	**Requires strong guidance**
Planning	The student participates in planning by making valuable proposals and making reasoned decisions.	The student participates in decision-making but does not make important proposals.	The student does not participate in goal planning but accepts the group's decision.	The student does not participate in planning goals and boycotts decisions made by peers.
Time	The student estimates the time required by the development steps of a product or task and adapts to possible changes and unforeseen events.	The student estimates the time required by the development steps of a product or task but has difficulty adapting to possible changes and unforeseen events.	The student does some scheduling and adjusts the task to the given schedule.	The student does not estimate the time needed for the task and does not get the task done in time.
Material	The student provides the necessary material to be able to work well in class and is concerned that this material is of good quality.	The student provides some material to be able to work in class, but the quality of the material is not very high.	The student provides material late or with errors that hinder the work.	The student does not provide enough material to be able to do the job properly.
Tranquillity	The student remains calm in the face of difficulties when performing a sequence of activities and finds alternative solutions.	The student has difficulties in keeping calm in the face of difficulties but ends up finding solutions.	The student keeps calm in the face of difficulties and asks for help or examples.	The student has difficulties in keeping calm in the face of difficulties as well as asking for help or accepting the given help.
Motivation	The student maintains motivation towards the activity from the first moment to the last.	The student starts the activity highly motivated but has difficulties in maintaining motivation throughout the course of the activity.	The student needs support from a group of peers or teachers to maintain motivation towards the activity.	The student's low motivation causes problems in the group or in the activity.
Help & respect	The student encourages other group members and recognizes the work they do. The student respects the limitations of others and helps them to improve and learn from them.	The student takes advantage of the potentialities of others and respects their limitations. However, the student cares more about getting the job done than helping others.	The student has difficulties in recognizing the potential of peers and becomes angry at their limitations. The student tends more to individual work than to group work.	The student does not recognize his/her groupmates and ends up working for her/himself instead of the group.
Delegation	The student accepts criticism from peers and corrects her/his actions. The student is aware of peers who can improve the work and delegates work to them.	The student hears criticism from peers but finds it hard to change his/her actions. The student delegates to peers what she/he knows they can do better.	The student gets angry with the criticism but ends up rectifying things. The student finds it difficult to delegate work to peers but manages to delegate some things.	The student does not accept criticism and does not know how to delegate. The students want to do everything because she/he does not trust her/his teammates.
Friendliness	If the student does not agree with what a peer says or does, he/she will be able to discuss the matter in proper and respectful language and suggest alternative solutions.	The student is able to discuss disagreements with the peers but does not suggest alternative solutions or may use an inappropriate way to express the matter.	If the student does not agree with what a peer says or does, he or she gets angry and does not want to talk about alternative solutions. The student gives up and does not look for alternative solutions.	If the student does not agree with what a peer says or does, he/she gets angry and provokes conflicts instead of looking for alternative solutions.

Students' Self-Assessment Tool

Part of the assessment tool is a student self-assessment. This gives the teacher a picture of how students perceive their own capabilities. The teacher is able to follow the development of the students in different areas. The competencies measured in the tool are transversal and development happens at the individual level over time with supportive exercises. Therefore, it is suggested to repeat the students' self-assessment once or twice a year to see the student's development. When students identify their own skills, it is possible for them to guide others as well (e.g. Slišāne et al., 2021). Additionally, peer learning has been found to be good in entrepreneurship education (e.g. Xu et al., 2021).

In the student self-assessment tool there is a list of statements dealing with the students' general feelings about themselves, their future visions and team working. The idea of the tool is that the students indicate the degree of agreement with each item. Students receive positive and encouraging feedback after answering the tool with ideas of how they can develop further with the mentioned EntreComp sub-competences. The feedback of the assessment tool is based on the scoring of the student's answers. The scale in the assessment is: Yes (2 points); Yes, to a certain degree (1 point) and No, not at all (0 points).

Table 6. Student self-assessment rubric

Vision (Ideas & opportunities)	Self-awareness and self-efficacy (Resources)	Working with others (Into action)
I have an idea of a dream job. I have a plan of how to achieve my dream job. I can easily come up with new ideas. I can easily get others excited about my ideas. I can imagine future visions. I can convince others that my vision is important.	I am aware of my strengths. I know how to improve my weaknesses. It is easy for me to stick to my goals. I can recognize everyday problems that need to be solved. When I am stuck with a problem I, can find a solution. If I try hard enough, I can achieve my goals.	When I am working in a team, I exchange information and ideas with other team members. When I am working in a team, I acknowledge others' ideas. When I am working in a team, I adhere to commonly agreed rules and schedules. I can lead my team finding new participatory solutions. I can build up a team to solve a specific problem. When I have a problem, I know whom to contact. I can solve conflicts in teams.

In this chapter a practical case is introduced related to the development of an assessment rubric. This is a rubric and an assessment tool linked to the assessment of the development of entrepreneurial competencies. In several curriculum texts, the development of entrepreneurial competence is described as a broad-based learning entity instead of the actual subject (e.g. Bourgeois et al., 2016; Finnish Ministry of Education and culture 2009 and 2016). This poses challenges for assessing competence development. There is a lack of clarity on whether and how it should be assessed. In addition, the concept of competence already involves contradictions and complexity. As mentioned earlier in this chapter, the definitions of the concept of entrepreneurship competence are somewhat confusing and overlapping (e.g. Bird 1995; Bridge et al. 2003, 37-39; Fastré and Van Gils 2007).

Based on the feedback received from teachers during the process, the use of this type of rubric in assessment provides added value to the teacher. However, as the rubric development process happened to occur during the covid-19 pandemic, there were challenges in committing teachers and students to

the development process. Although the training course and piloting were successfully completed, there is always a risk that the use of the assessment tool on a regular basis is forgotten. In order to make the assessment of these competencies more general, it would be important for them to be included in the national curriculum of each country and further in the curriculums and yearly plans of schools.

During the process teachers were introduced to concrete methods for developing entrepreneurial competencies. The rubric clearly made what was to be assessed visible when assessing entrepreneurial competence. However, there is still a long way to go before a rubric exists for teachers to consciously use the rubric as part of their daily teaching. On the other hand, students should also understand what the concepts in the rubric mean. In other words, teachers and students need to work together to build an understanding of what entrepreneurial competence means, how it can be developed and how it can be assessed. This forms the basis for the student's self-assessment and the assessment carried out by the teacher.

The participatory process changed the perspective on developing the rubric. Teachers chose sub-competencies, and had the opportunity to develop criteria for the competencies so that they understood them and could commit to them. On the other hand, creating the tool together with students significantly affected the usability of the tool. The language should be such that it is clear to young users. The implementation of the tool should be such that self-assessment can be done using a natural tool for the students. The participatory development process increased the understanding of the different cultural factors that influence the implementation of the assessment. It should be noted that the assessment should be appropriate for different school environments and cultures, even if the competencies to be developed are universal. In addition, understanding the concepts used in the assessment rubrics is culturally binding and this must be taken into account when developing this type of rubric.

The rubrics presented in this chapter are based on teachers' choices and the way they understand entrepreneurship competence. Teachers implement entrepreneurship education, so their interpretations matter. School management should pay attention to the fact that the teachers have the time, opportunities, and space to build an understanding of entrepreneurial competence together with other teachers. As noted in a previous research (Sommarström et al., 2021), teacher autonomy is important in understanding entrepreneurship education and, on the other hand, school management affects how a teacher is able to implement entrepreneurship education (Hämäläinen et al., 2018; Ruskovaara et al. 2016).

The rubrics presented in this chapter have some major weaknesses that need to be pointed out. These rubrics offer an opportunity for further development. Teachers have clearly had challenges in articulating the assessment criteria (vision, self-awareness and self-efficacy and recourses). It is possible that more time would have been needed for the teachers to work on the rubric and to find different ways and working methods to create a common understanding of the sub-competencies. Cultural differences, as well as the teachers' own background, and language influenced the teachers' interpretations of the sub-competencies and their assessment criteria. However, this presented model provides a good basis for further work and, above all, emphasizes the importance of involving different user groups in the work.

During the process, activities were created to enable the teacher to promote students' entrepreneurial competence in face-to-face educational situations. These were practical entrepreneurial activities (i.e. reflective exercises, card games and problem-solving techniques). These activities support teachers in planning and implementing entrepreneurship education in the classroom. They also enhance the use of the assessment rubric. In our experience, the use of the rubric requires practice, and various active approaches can promote its use. There was a training path for teachers in a specific training course. The training included strategies for teaching entrepreneurial competence, strategies for the teaching and use

of a digital tool for the assessment and student self-assessment activities. After the training, teachers were asked about their perceptions of the learning outcomes the training offered. Approximately 78 percent of the teachers attending the training were totally satisfied with the learning outcomes (see e.g. Castañeda et al., 2021). In addition, about 70 percent of the teachers who participated in the coaching felt that their ability to develop students' entrepreneurial skills improved as a result of the training.

FUTURE RESEARCH DIRECTIONS

This chapter has presented one example of how to assess the development of learners' entrepreneurial competence. As entrepreneurship education has extended to all levels of education, teachers on every school level should have the tools and skills to assess the development of learners' entrepreneurship competence. Based on the experiences gained from the development process presented in this chapter, there is a need for more profound research concentrating on how entrepreneurship competence should be assessed. The development of entrepreneurship competence rubrics would help to verbalize age-appropriate entrepreneurial competences. Rubrics and tools for assessment needs to be further developed and evaluated.

In addition, the tool and rubric presented in this chapter only address some of the general entrepreneurial competencies of EntreComp framework. It is worth noting that teachers selected the three competencies with which they found it most difficult to assess student development. Other sub-competencies also require their own assessment framework. Hence, for further studies there is a great need to build an assessment rubric for all the entrepreneurial competences included in EntreComp framework. Based on our experience, it should be done according to the principles of co-development, with teachers and students (rubric users) participating in the creation of assessment models.

CONCLUSION

The main purpose of the chapter was to illustrate the construction of an entrepreneurial competence assessment rubric. This development activity originates from an ambition to train and strengthen teacher's competence in integrating entrepreneurship into their daily teaching. The goal in the end is for teachers to be able to evaluate and assess entrepreneurial abilities and competencies of learners in primary and basic school (students aged 7-18years).

Through the improvement of the assessment skills of teachers, the assessment tool and rubrics presented in this chapter help students in acquiring/evaluating entrepreneurial competence, allowing them to develop transversal skills such as critical thinking, creativity, and self-perception, and introducing them to the professional world and promoting their active citizenship. So, the general objective is to produce an impact on the school system, in terms of improvement of the education quality and evaluation and self-evaluation of entrepreneurial skills, through the update of teachers' skills.

In addition, in several curriculum texts, the development of entrepreneurial skills is described as a broad-based learning entity instead of an actual subject (e.g. Finnish National Agency for Education (2016). Bourgeois et al., 2016). This poses challenges in assessing competence development. The aim is to develop entrepreneurial skills, but there is a lack of clarity on whether and how they should be assessed. In addition, the concept of entrepreneurial competence already involves contradictions and

complexity. In addition, when it comes to assessing the abilities of young people, it mixes the characteristics developed by the school as well as the personality traits of the individual. The measurement needs testing and further development.

Policies fostering entrepreneurial competencies in schools do not directly lead to implementing entrepreneurial methods in practice. Objectives should be clearly stated in schools' curriculums and supported by school management. One of the means supporting teachers in their implementation of methods that develop entrepreneurship competence is to provide continuing training for teachers.

In conclusion, despite extensive and diverse research related to entrepreneurship competencies (e.g. Komarkova et al. 2015; Lilleväli, 2017; Tsakiridou & Stergiou, 2014) Venesaar et al., 2021), the development of tools suitable for assessing the development of entrepreneurial competencies has been surprisingly limited. In our view, a rubric is a model that can be used to strengthen the teacher's competence in assessment and in understanding entrepreneurial competence in general. However, evaluation cannot be implemented as a separate measure, but must be linked to the development of the institution's activities. The role of school management is central to how the assessment is approached in the first place and what opportunities the teacher has to carry out the assessment. This requires training, understanding, and collaboration within the school.

REFERENCES

Almeida, F., & Buzady, Z. (2019). Assessment of entrepreneurship competencies through the use of FLIGBY. *Digital Education Review*, 151-169. doi:10.1344/der.2019.35.151-169

Andrade, H. (2000). Using rubrics to promote thinking and learning. *Educational Leadership*, *57*(5), 13–18.

Andrade, H., & Du, Y. (2005). Student perspectives on rubric-referenced assessment. *Practical Assessment, Research & Evaluation*, *10*(1), 3. doi:10.7275/g367-ye94

Athayde, R. (2009). Measuring enterprise potential in young people. *Entrepreneurship Theory and Practice*, *33*(2), 481–500. doi:10.1111/j.1540-6520.2009.00300.x

Babatunde, S., & El-Gohary, H. (2019). Necessity of mentoring in entrepreneurship education: Reflection by practitioners. *Journal of Professional Issues in Engineering Education and Practice*, *145*(1), 02518007. doi:10.1061/(ASCE)EI.1943-5541.0000399

Babatunde, S., El-Gohary, H., & Edwards, D. (2021). Assessment methods in entrepreneurship education, challenges and opportunities in developed and developing nations: A comparative study of Nigeria and England. *Education + Training*, *63*(7/8), 1092–1113. doi:10.1108/ET-12-2020-0368

Bacigalupo, M., Kampylis, P., Punie, Y., & Van den Brande, G. (2016). EntreComp: The entrepreneurship competence framework. *Luxembourg. Publication Office of the European Union*, *10*, 593884.

Bandura, A. (1977). Self-efficacy: Toward a unifying theory of behavioral change. *Psychological Review*, *84*(2), 191–215. doi:10.1037/0033-295X.84.2.191 PMID:847061

Baron, R. A. (2006). Opportunity recognition as pattern recognition: How entrepreneurs "connect the dots" to identify new business opportunities. *The Academy of Management Perspectives*, *20*(1), 104–119. doi:10.5465/amp.2006.19873412

Bergevoet, R. H., & Woerkum, C. V. (2006). Improving the entrepreneurial competencies of Dutch dairy farmers through the use of study groups. *Journal of Agricultural Education and Extension*, *12*(1), 25–39. doi:10.1080/13892240600740852

Beverland, M., & Lindgreen, A. (2010). What Makes a Good Case Study? A Positivist Review of Qualitative Case Research Published in Industrial Marketing Management, 1971–2006. *Industrial Marketing Management*, *39*(1), 56–63. doi:10.1016/j.indmarman.2008.09.005

Bird, B. (1995). Towards a Theory of Entrepreneurial Competency. *Advances in Entrepreneurship, Firm, Emergence, and Growth*, *2*(1), 51–72. doi:10.1108/S1074-754020190000021011

Blenker, P., Elmholdt, S. T., Frederiksen, S. H., Korsgaard, S., & Wagner, K. (2014). Methods in entrepreneurship education research: A review and integrative framework. *Education + Training*, *56*(8/9), 697–715. doi:10.1108/ET-06-2014-0066

Bourgeois, A., Balcon, M. P., & Riiheläinen, J. M. (2016). *Entrepreneurship Education at School in Europe. Eurydice 2016 Report.* Education, Audiovisual and Culture Executive Agency, European Commission. https://eacea.ec.europa.eu/national-policies/Bourgeois

Boyd, N. G., & Vozikis, G. S. (1994). The influence of self-efficacy on the development of entrepreneurial intentions and actions. *Entrepreneurship Theory and Practice*, *18*(4), 63–77. doi:10.1177/104225879401800404

Bresciani, M. J., Zelna, C. L., & Anderson, J. A. (2004). Criteria and rubrics. In *Assessing Student Learning and Development: A Handbook for Practitioners* (pp. 29–37). National Association of Student Personnel Administrators.

Bridge, S., O'Neill, K., & Cromie, S. (2003). *Understanding enterprise. Entrepreneurship and Small Business*. Palgrave McMillan.

Brophy, M. and Kiely, T. (2002). Competencies: a new sector. *Journal of European Industrial Training, 26* (2/3/4), 165-176. doi:10.1108/03090590210422049

Castañeda, L., Talborn Björkvi, S., Tilly, A., Minin, D., Hernández, I., & Hämäläinen, M. (2021). *Aprendizaje conectado como práctica sistémica para procesos de desarrollo profesional docente: un estudio de caso basado en la combinación de estrategias*.

Cotoi, E., Bodoasca, T., Catana, L., & Cotoi, I. (2011). Entrepreneurship European development strategy in the field of education. *Procedia: Social and Behavioral Sciences*, *15*, 3490–3494. doi:10.1016/j.sbspro.2011.04.323

Creswell, J. W. (1998). *Qualitative inquiry and research design: Choosing among five traditions*. Sage.

Cuenca, L., Alarcón Valero, F., Boza, A., Fernández Diego, M., Ruiz Font, L., Gordo Monzó, M. L., & Alemany Díaz, M. D. M. (2016). Rubric to Assess the Competence OF Innovatoon, Creativity and Entrepreneurship in Bachelor Degree. *Brazilian Journal of Operations & Production Management*, *13*(1), 118–123. doi:10.14488/BJOPM.2016.v13.n1.a14

de Lourdes Cárcamo-Solís, M., del Pilar Arroyo-López, M., del Carmen Alvarez-Castañón, L., & García-López, E. (2017). Developing entrepreneurship in primary schools. The Mexican experience of "My first enterprise: Entrepreneurship by playing". *Teaching and Teacher Education*, *64*, 291–304. doi:10.1016/j.tate.2017.02.013

European Commission. (2012). *Vocational education and training for better skills, growth and jobs*. EC. https://eur-lex.europa.eu/legal-content/EN/TXT/PDF/?uri=CELEX:52012SC0375&from=EN

European Commission. (2013). *Entrepreneurship 2020 action plan: Reigniting the entrepreneurial spirit in Europe*. European Commission.

Farquhar, J. D. (2012). *Case Study Research for Business*. SAGE. doi:10.4135/9781446287910

Fastré, G., & Van Gils, A. (2007). Competence development in Entrepreneurship. The Role of University Education. In M.K. McCuddy (pp. 385-398) The challenges of educating people to lead in a challenging world. Springer, Dordrecht. doi:10.1007/978-1-4020-5612-3_19

Ferreras-Garcia, R., Hernández-Lara, A. B., & Serradell-López, E. (2019). Entrepreneurial competences in a higher education business plan course. *Education + Training*, *61*(7/8), 850–869. doi:10.1108/ET-04-2018-0090

Fiet, J. (2001). The Theoretical Side of Teaching Entrepreneurship. *Journal of Business Venturing*, *16*(1), 1–24. doi:10.1016/S0883-9026(99)00041-5

Figueiredo-Nery, M., N., A., & Figueiredo, P. (2008). Forming entrepreneurial mindsets? Preliminary evidence of teaching practices from primary schools in a developing area in south America. *Journal of Technology Management & Innovation*, *3*(2), 1–17.

Finnish Ministry of Education. (2009). *Guidelines for entrepreneurship education*. FME. https://julkaisut.valtioneuvosto.fi/handle/10024/78871

Finnish Ministry of Education and Culture. (2016). *Finnish education in a nutshell*. FMEC. https://www.oph.fi/en/statistics-and-publications/publications/finnish-education-nutshell

Finnish National Agency for Education. (2016). National Core Curriculum for Basic Education. [Helsinki.]. *Publications*, *2016*, 5.

Garavan, T. N., & McGuire, D. (2001). Competencies and workplace learning: Some reflections on the rhetoric and the reality. *Journal of Workplace Learning*, *13*(4), 144–164. doi:10.1108/13665620110391097

Gibb, A. (2002). In pursuit of a new 'enterprise' and 'entrepreneurship' paradigm for learning: Creative destruction, new values, new ways of doing things and new combinations of knowledge. *International Journal of Management Reviews*, *4*(3), 233–269. doi:10.1111/1468-2370.00086

Goodrich, H. (1997). Understanding rubrics. *Educational Leadership*, *54*(4), 14–17.

Gustafsson-Pesonen, A., & Remes, L. (2012). Evaluation of entrepreneurial development coaching: Changing the Teachers' thinking and action on entrepreneurship. *Annals of Innovation amp; Entrepreneurship*, *3*(1), 17211. doi:10.3402/aie.v3i0.17292

Hämäläinen, M., Ruskovaara, E., & Pihkala, T. (2018). Principals promoting entrepreneurship education: The relationships between development activities and school practises. *Journal of Entrepreneurship Education*, *21*(2), 1–19.

Hamel, J., Dufour, S., & Fortin, D. (1993). *Case Study Methods*. SAGE. doi:10.4135/9781412983587

Henry, C., Hill, F., & Leitch, C. (2005). Entrepreneurship education and training: Can entrepreneurship be taught? Part I. *Education + Training*, *47*(2), 98–111. doi:10.1108/00400910510586524

Henry, C., & Lewis, K. (2018). A review of entrepreneurship education research: Exploring the contribution of the Education + Training special issues. *Education + Training*, *60*(3), 263–286. doi:10.1108/ET-12-2017-0189

Jerusalem, M., & Schwarzer, R. (1995). *General Self-Efficacy Scale—Revised—English Version (Gse-R)* [Database record]. APA PsycTests. doi:10.1037/t18916-000

Johansen, V., & Schanke, T. (2013). Entrepreneurship education in secondary education and training. *Scandinavian Journal of Educational Research*, *57*(4), 357–368. doi:10.1080/00313831.2012.656280

Jones, B., & Iredale, N. (2010). Enterprise education as pedagogy. *Education + Training*, *1*(52), 7–19. doi:10.1108/00400911011017654

Komarkova, I., Conrads, J., & Collado, A. (2015). Entrepreneurship Competence: An Overview of Existing Concepts. *Policies and Initiatives. depth case study report*. https://citeseerx.ist.psu.edu/viewdoc/download?doi=10.1.1.874.7140&rep=rep1&type=pdf

Kujala, S. (2003). User involvement: A review of the benefits and challenges. *Behaviour & Information Technology*, *22*(1), 1–16. doi:10.1080/01449290301782

Kyndt, E., & Baert, H. (2015). Entrepreneurial competencies: Assessment and predictive value for entrepreneurship. *Journal of Vocational Behavior*, *90*, 13–25. doi:10.1016/j.jvb.2015.07.002

Lackéus, M., & Sävetun, C. (2019). Assessing the impact of enterprise education in three leading Swedish compulsory schools. *Journal of Small Business Management*, *57*(sup1), 33–59. doi:10.1111/jsbm.12497

Lans, T., & Gulikers, J. (2010). Assessing entrepreneurial competence in entrepreneurship education and training. Handbook of research in entrepreneurship education, 3, 54-67.

Le Deist-Delamare, F. D., & Winterton, J. (2005). What is competence? *Human Resource Development International*, *8*(1), 27–46. doi:10.1080/1367886042000338227

Lilleväli, U., & Täks, M. (2017). *Competence models as a tool for conceptualizing the systematic process of entrepreneurship competence development*. Education Research International. doi:10.1155/2017/5160863

Man, T. W. Y., Lau, T., & Chan, K. F. (2002). The competitiveness of small and medium enterprises: A conceptualisation with focus on entrepreneurial competencies. *Journal of Business Venturing*, *17*(2), 123–142. doi:10.1016/S0883-9026(00)00058-6

Markman, G. D., Baron, R. A., & Balkin, D. B. (2005). Are perseverance and self-efficacy costless? Assessing entrepreneurs' regretful thinking. *Journal of Organizational Behavior: The International Journal of Industrial. Journal of Organizational Behavior*, *26*(1), 1–19. doi:10.1002/job.305

McGee, J. E., Peterson, M., Mueller, S. L., & Sequira, M. J. (2009). Entrepreneurial self-efficacy: Refining the measure. *Entrepreneurship Theory and Practice*, *33*(4), 965–988. doi:10.1111/j.1540-6520.2009.00304.x

Mitchelmore, S., & Rowley, J. (2010). Entrepreneurial competencies: A literature review and development agenda. *International Journal of Entrepreneurial Behaviour & Research*, *16*(2), 92–111. doi:10.1108/13552551011026995

Moskal, B. M. (2003). Recommendations for developing classroom performance assessments and scoring rubrics. Practical Assessment, Research & Evaluation, 8(1), 14. ttps:// doi:10.7275/jz85-rj16

Obschonka, M. (2016). Adolescent pathways to entrepreneurship. *Child Development Perspectives*, *10*(3), 196–201. doi:10.1111/cdep.12185

Pittaway, L., & Edwards, C. (2012). Assessment: Examining practice in entrepreneurship education. *Education + Training*, *54*(8/9), 778–800. doi:10.1108/00400911211274882

Popham, J. W. (1997). What's wrong-and what's right-with rubrics. *Educational Leadership*, *55*(2), 72–75.

Reddy, Y. M., & Andrade, H. (2010). A review of rubric use in higher education. *Assessment & Evaluation in Higher Education*, *35*(4), 435–448. doi:10.1080/02602930902862859

Robinson, S., & Stubberud, H. A. (2014). Teaching creativity, team work and other soft skills for entrepreneurship. *Journal of Entrepreneurship Education*, *17*(2), 186.

Ruskovaara, E., Hämäläinen, M., & Pihkala, T. (2016). HEAD teachers managing entrepreneurship education–Empirical evidence from general education. *Teaching and Teacher Education*, *55*, 155–164. doi:10.1016/j.tate.2016.01.004

Ruskovaara, E., & Pihkala, T. (2013). Teachers implementing entrepreneurship education: Classroom practices. *Education + Training*, *55*(2), 204–216. doi:10.1108/00400911311304832

Schelfhout, W., Bruggemana, K., & De Maeyerb, S. (2016). Evaluation of entrepreneurial competence through scaled behavioural indicators: Validation of an instrument. *Studies in Educational Evaluation*, *51*, 29–41. doi:10.1016/j.stueduc.2016.09.001

Scholz, U., Doña, B. G., Sud, S., & Schwarzer, R. (2002). Is general self-efficacy a universal construct? Psychometric findings from 25 countries. *European Journal of Psychological Assessment*, *18*(3), 242. doi:10.1027//1015-5759.18.3.242

Schumpeter, J. A. (1982). *The theory of economic development: An inquiry into profits, capital, credit, interest, and the business cycle. Harward Economic Studies 46*. Harvard University Press.

Seikkula-Leino, J. (2011). The implementation of entrepreneurship education through curriculum reform in Finnish comprehensive schools. *Journal of Curriculum Studies*, *43*(1), 69–85. doi:10.1080/00220270903544685

Short, J. C., Ketchen, D. J. Jr, Shook, C. L., & Ireland, R. D. (2010). The concept of "opportunity" in entrepreneurship research: Past accomplishments and future challenges. *Journal of Management*, *36*(1), 40–65. doi:10.1177/0149206309342746

Silveyra, G., Herrero, Á., & Pérez, A. (2021). Model of teachable entrepreneurship competencies (M-TEC): Scale development. *International Journal of Management Education*, *19*(1), 100392. doi:10.1016/j.ijme.2020.100392

Sommarström, K., Oikkonen, E., & Pihkala, T. (2021). The school and the teacher autonomy in the implementing process of entrepreneurship education curricula. *Education Sciences*, *11*(5), 215. doi:10.3390/educsci11050215

Stevens, D.D. & Levi, A. J. (2005). Leveling the field: Using rubrics to achieve greater equity in teaching and assessment. *Essays on Teaching Excellence, Professional and Organizational Development Network in Higher Education, 17* (1).

Tahahashi, K., & Kiyosumi, M. (2021). Development of Rubric for enhancing Sensemaking among Team members in PBL in Entrepreneurship Education. In *International Symposium on Affective Science and Engineering ISASE2021*. Japan Society of Kansei Engineering. 10.5057/isase.2021-C000004

Tsakiridou, H., & Stergiou, K. (2014). Entrepreneurial competences and entrepreneurial intentions of students in primary education. *International Journal of Humanities Social Sciences and Education*, *1*(9), 106–117.

Ustav, S., & Venesaar, U. (2018). Bridging metacompetencies and entrepreneurship education. *Education + Training*, *60*(7/8), 674–695. doi:10.1108/ET-08-2017-0117

Venesaar, U., Malleus, E., Arro, G., & Toding, M. (2021). Entrepreneurship Competence Model for Supporting Learners Development at All Educational Levels. *Administrative Sciences*, *12*(1), 2. doi:10.3390/admsci12010002

Verschuren, P. (2003). Case study as a research strategy: Some ambiguities and opportunities. *International Journal of Social Research Methodology*, *6*(2), 121–139. doi:10.1080/13645570110106154

Woodruffe, C. (1993). What Is Meant by a Competency? *Leadership and Organization Development Journal*, *14*(1), 29–36. doi:10.1108/eb053651

Yin, R. K. (2012). Case study methods. In H. Cooper, P. M. Camic, D. L. Long, A. T. Panter, D. Rindskopf, & K. J. Sher (pp. 141–155.), APA handbook of research methods in psychology, Vol. 2. Research designs: Quantitative, qualitative, neuropsychological, and biological. American Psychological Association. doi:10.1037/13620-009

ADDITIONAL READING

Hämäläinen, M. (2023). *Principals managing entrepreneurship education in schools.*

Hämäläinen, M., Joensuu-Salo, S., Peltonen, K., & Raappana, A. (2022). HEI teacher perceptions of entrepreneurship education: The role of teachers' entrepreneurial backgrounds and HEI managerial support. In *Strategies for the Creation and Maintenance of Entrepreneurial Universities* (pp. 114–141). IGI Global. doi:10.4018/978-1-7998-7456-0.ch006

Joensuu-Salo, S., Peltonen, K., Hämäläinen, M., Oikkonen, E., & Raappana, A. (2021). Entrepreneurial teachers do make a difference – Or do they? *Industry and Higher Education*, *35*(4), 536–546. doi:10.1177/0950422220983236

Minna, H., Elena, R., & Timo, P. (2018). Principals promoting entrepreneurship education: The relationships between development activities and school practises. *Journal of Entrepreneurship Education*, *21*(2), 1–19.

Oksanen, L., Oikkonen, E., & Pihkala, T. (2023). Adopting Entrepreneurship Education—Teachers¢ Professional Development. *Entrepreneurship Education and Pedagogy*, *6*(2), 276–298. doi:10.1177/25151274221091698

Raappana, A., Pihkala, T., & Kuru, P. (2021). Finnish ninth graders' perception of their entrepreneurial selves: general overview of the 2020 survey round.

Ruskovaara, E., & Pihkala, T. (2013). Teachers implementing entrepreneurship education: Classroom practices. *Education + Training*, *55*(2), 204–216. doi:10.1108/00400911311304832

Ruskovaara, E., Pihkala, T., Seikkula-Leino, J., & Rytkölä, T. (2015). Creating a measurement tool for entrepreneurship education: A participatory development approach. *Developing, shaping and growing entrepreneurship*, 40-57.

Sommarström, K., Ruskovaara, E., & Pihkala, T. (2017). Company visits as an opportunity for entrepreneurial learning. *Journal for International Business and Entrepreneurship Development*, *10*(3), 298–315. doi:10.1504/JIBED.2017.085505

KEY TERMS AND DEFINITIONS

Assessment Rubric: A tool used by educators to evaluate and provide feedback on students' performance, understanding, and skills. It is a systematic way to assess and communicate the expectations and criteria for successful completion of a task or assignment. Rubrics in learning are particularly helpful for projects, essays, presentations, and other complex tasks where subjective judgment may be involved.

EntreComp: The European Entrepreneurship Competence Framework has developed by European Commission as a reference framework to explain what is meant by an entrepreneurial mindset.

Entrepreneurial Competence: The set of skills, knowledge, and abilities that enable individuals to be entrepreneurial and create financial, cultural or social value for others.

Entrepreneurship Education: Learning through, for, and about entrepreneurship, taking place in an entrepreneurial learning environment, and being about dealing with, creating and enjoying uncertainty and complexity. The aim is to strengthen students' personal and generic entrepreneurial competences.

Entrepreneurship: The process of creating, developing, and managing a new business venture or startup.

Online Tool: Tool is available on website by using username and password. All data is Downloaded or Uploaded.

Self-Assessment: Process of students evaluating their own work, skills, and understanding of concepts. It is a critical aspect of the learning process, as it encourages individuals to reflect on their performance and identify areas of strength and weakness.

Chapter 4
Learning How to Become a Teacher Researcher:
Using Rubrics to Support Evidence-Informed, Research-Based Practice

Emma O. Brien

Department of Language and Literacy Education, Faculty of Education, Mary Immaculate College, Limerick, Ireland

Josephine Brady

Department of Language and Literacy Education, Faculty of Education, Mary Immaculate College, Limerick, Ireland

T. J. Ó Ceallaigh

School of Education, University College Cork, Cork, Ireland

Katharine Babbitt

Department of Language and Literacy Education, Faculty of Education, Mary Immaculate College, Limerick, Ireland

Andrea Brosnan

Department of Language and Literacy Education, Faculty of Education, Mary Immaculate College, Limerick, Ireland

Emma Byrne

Department of Language and Literacy Education, Faculty of Education, Mary Immaculate College, Limerick, Ireland

Erin Byrne

Department of Language and Literacy Education, Faculty of Education, Mary Immaculate College, Limerick, Ireland

Rebecca Curtin

Department of Language and Literacy Education, Faculty of Education, Mary Immaculate College, Limerick, Ireland

Lisa Gaffney

Department of Language and Literacy Education, Faculty of Education, Mary Immaculate College, Limerick, Ireland

Karen O'Callaghan

Department of Language and Literacy Education, Faculty of Education, Mary Immaculate College, Limerick, Ireland

ABSTRACT

Developing the teacher as a reflective practitioner has become a core facet of Irish teacher education, but large gap exists between theory and practice. Research illustrates the positive value of rubrics in terms of student self-reflection and self-regulation. However, few studies explore the use of assessment rubrics

within the broad context of the supervisory relationship. Drawing on Drytons extended supervisory working alliance the authors explored how rubrics can foster student-supervisor relationships during the research process. This study adopted a collaborative autoethnographic (CAE) methodology which enabled the researchers to authentically capture the student and faculty perspectives. It was found that rubrics provided transparency, identified expectations and a language for students to express and inter- rogate their work. They provided a metadialogue to enable students to take ownership of the feedback process questioning and initiating discussion with the supervisor. This supported the bonding process, shaping conversations and providing the student with context for the feedback.

INTRODUCTION

Calls for research-based initial teacher education (ITE) programmes have increased over the past decade. Many novice teachers struggle to transform ITE knowledge into effective classroom practice (Doyle, 2006; Wanzare, 2007) which calls for a deeper understanding of teacher transition from theory to practice (OECD, 2021). Student teachers must be active participants in the research process during ITE (Cochran-Smith & Zeichner, 2005; Darling-Hammond, 2017; Munthe & Rogne, 2015). It is not surprising that the development of teacher researchers is often discussed within the context of ITE reform and development (Afdal & Speres, 2018; Darling-Hammond, 2016, 2017). The emancipatory potential of teacher research is also recognised in ITE in the Republic of Ireland. For example, Céim, i.e. Standards in initial teacher education, states that by the end of the ITE programme, the student teacher will be able to "conduct and apply relevant research as appropriate to his/her teaching context, identifying, critically analysing and integrating new knowledge regarding curriculum, pedagogy and assessment into their practice" (pg.23). Developing the teacher as 'reflective practitioner and 'researcher' has become a core facet of Irish teacher education (Teaching Council, 2017, p.14). However research is interactive and incremental, learning occurs through engagement in student-supervisor dialog, self-dialog/reflection and the research process. This requires a process orientated approach to assessment rather than the traditional product and standards based model. Therefore assessment AS learning is key with reflection, self-assessment and dialogical feedback all necessary components to support the research and assessment process. It is the authors view that the integration of formative and summative rubrics into the supervision process can support the development of teachers as researcher. Rubrics can provide a lexis to enable the student-supervisor to articulate feedback, communicate and engage in dialog and also enable the students to engage in critical reflection regarding the research process and their own identity and role. This chapter discusses the role of the supervisor-student relationship in teacher research, the potential for rubrics to support the process and outlines a collaborative autoethnographic study which the authors engaged in to explore how formative rubrics can support the development of teacher-researchers within the research-based teacher education movement. In particular, it will consider the role of formative rubrics in fostering student-supervisor relationships during the ITE research process.

THE TEACHER RESEARCHER: TRADITIONAL APPROACHES TO SUPERVISION

Within the educational field, there is a strong tradition of teachers engaging in research (see Elliot, 1991; Lytle & Cochran-Smith, 1994; Zeichner and Noffke, 2001 for the history of teacher research); a tradition helped greatly by Giroux's (1988) scholarship on 'teachers as intellectuals' and Cochran-Smith &Lytle (1992, 1999) prioritisation of the 'teachers as researchers' discourse.

Put simply: "the key to the whole approach is the role of the teacher as researcher" (1975:141); a teacher who is "studying themselves" (ibid) as well as applying their tentative understandings in their own contexts. Stenhouse envisages teachers as self-reflective, self- monitoring 'artists':

A teacher lays the foundation of his capacity for research by developing self-monitoring strategies. The effect is not unlike that of making the transition from amateur to professional actor. Through self-monitoring the teacher becomes a conscious artist. (Stenhouse, 1985: 15)

However, whilst Stenhouse's process model is now universally well-regarded, it was not accepted by mainstream education at the time, largely because, as Stenhouse himself acknowledged, "the process model is essentially a critical model, not a marking model" (1975: 95).

In today's parlance, Stenhouse's process model is akin to continuous self-assessment as he saw the process of thinking and seeking to understand something at a deeper level as ongoing and as far more important than the content itself.

The superficialities of the disciplines may be taught by pure instruction but the capacity to think within the disciplines can only be taught by inquiry. (1975: 38)

These skills cannot be taught and involve dialog, critical self-assessment and reflection. Therefore from Stenhouse's point of view, the role of supervisors was to help teachers enact their 'teacher-researcher' identities, develop their understandings and engage in transformative dialog.

Cohran-Smith & Lytle (1993) coined the term 'inquiry as stance' where they argued for reflective and agentic teachers who adopt the overarching stance or mindset of inquiring learners. They said:

Inquiry as stance is neither a top-down nor a bottom-up theory of action, but an organic and democratic one that positions practitioners' knowledge, practices, and their interactions with students and other stakeholders at the center of educational transformation. (2009: 123-124).

The key message was that teaching was no longer to be seen solely in terms of 'doing' but rather required an ongoing professional commitment to 'knowing, doing and being'. The relationship between teacher research and professional development was dynamic and personal, but according to Tafel and Fischer (1996) there were four identifiable elements: 1) caring relationships and dialogue- which relates to the supervisor-teacher relationships; 2) teacher agency and responsibility to the communities of learning; 3) knowledge creation through reflection and 4) inquiry based on their guiding principles and beliefs as teachers.

In particular the importance of the supervisory relationship is recognized (Mainhard et al., 2009; Parker-Jenkins, 2018) yet comparatively little is known about the intricacies of supervisory interactions and the connection with formative assessment processes, particularly at master's level.

Building on the work of Mouton (2001) and Gatfield (2005), Lee (2012) described five supervisory styles: functional – directing students through research goals and tasks; enculturation –into the academic discipline; critical thinking; emancipation and relationship development. Against the backdrop of Dysthe (2002) three static models of supervision: a teaching model, a partnership model and an apprenticeship model, Lee (2012) offers a far more nuanced and complex picture of supervisory relationships. He paints

a picture which allows for shifts over time and for the student to become a researcher in his or her own right (Mc Alphine, 2013; Bui, 2014). As Gurr (2001) recognised:

The supervisory style needs to be adjusted to a more hands-off approach to allow competent autonomy to be developed. (86-87)

De Kleijn et al (2015) describe this as 'adaptive research supervision' which accords with Brook et al.'s (2010: 665) view of postgraduate learning as 'pathmaking' which involves embarking on a "collaborative quest for learning that encompasses exploration of the unknown, discovering new understandings" through "a creative, richly textured, open-ended" process. Positive supervisory relationships are characterised by two forms of dialogue: transformative and explorative (Sarja and Janhonen's, 2009) Transformative dialogue engages both the supervisor and supervisee in self-reflection and disclosure as well as problem-solving discussions while exploratory dialogue is built upon the willingness of both participants to share their tacit knowledge and go beyond the traditional didactic supervisory relationship.

Positive supervisory relationships have been linked to student well –being, progression, timely completion and career development (Sverdlik et al, 2018; McAlphin & McKinnon, 2013; Lin, 2012). Using a concept borrowed from psychotherapy, Frischer & Larsson (2000) suggest that a 'working alliance' should be established to support collaboration and mitigate against a laissez-faire approach to supervision. Many scholars have also identified significant barriers such as conflicting expectations (Zhao et al, 2007), power differential (Moxham et al, 2013) and a sense of isolation from the wider research community (Pilbeam & Denyer, 2009). Barriers which have been described as 'perennial issues' which can only be overcome through proactive supervisory interactions. Traditionally rubrics have be associated with the assessment and marking component of the research process, to date no exploration has been conducted in the role of rubrics in supporting supervisor-student relationships and how they can enable transformative dialog and common language around research. The next section will explore existing research on the role of rubrics in teaching, learning and the supervision process.

RUBRICS AND THEIR ROLE IN TEACHING AND LEARNING AND SUPERVISION

Rubrics and Their Role in Teaching and Learning

Assessment forms an integral part of students teaching and learning experience and often drives what and how students engage with learning. (Boud and Falchikov 2006; Cullen and Harris 2009) Students often see assessment as an 'end game' This influenced by the traditional view of assessment which was often perceived as a means of measuring academic performance with the objective of grading students for the purposes of an award. However recently the discourse surrounding assessment has shifted significantly from assessment of learning to assessment as and for learning. (National Forum, 2016). Key to this shift is formative assessment and the role of relationships, dialogue and feedback. This supports students in self-assessing the quality of their work. Black and William also advocate that formative assessment supports the students to 1) recognize what is required to evidence their learning for the assessment, 2) consider the quality of work in relation to academic standards, and 3) have an understanding how to improve the quality of their work. (Nicol 2009; Sadler 2009). Students can then adapt their learning and assessments forms a key part of their learning environment. To successfully engage in the formative assessment and feedback process students need to be able to critically evaluate and judge work based

on academic standards. (Nicol 2009). Rubrics can be used as a means of defining such standards and supporting students to evaluate their work, particularly in the research supervision process

A rubric is traditionally defined as 'a document that articulates the expectations for an assignment by listing the criteria, or what counts, and describing levels of quality from excellent to poor' (Andrade and Du 2005). This aligns to the traditional view of assessment whereby they were used to measure academic performance. However, as the discourse of assessment as and for learning has emerged they have been widely adopted as an instructional tool. This has led to the emergence of a variety of ways of designing rubrics that often lecturers and supervisors are not aware of.

Rubrics can be generic or task specific (Jonasson and Svingby, 2007). Generic Rubrics may be based on the type of assessment e.g. a generic rubric for essays, presentations, literature reviews. Or based on institutional criteria e.g. to standardise assessment and feedback in an education institute there might be a decision to adopt a generic rubric. These may be useful for students to track their progress across similar assessments or programmes and benchmark their learning development.

Task or discipline specific are developed based on the learning outcomes and of the module, programme or course and the tasks associated with evidencing this as part of the assessment or which are aligned to the learning outcomes. Students often find task specific rubrics easier to interpret or use as they are context specific and the language is not as vague as those in generic rubrics (Bailey, 2009; Jonsson, 2014). In addition, the criteria in the rubric may evaluate a variety of dimensions for example discipline specific knowledge, soft skills and transferability to the professional context (Dwyer et al 2006)

Furthermore, rubrics can be designed to be analytic or holistic (Wolcott and Legg 1998, 71). Holistic rubrics consider the overall quality of the students work and how sections inter-relate rather than grading individual parts. Alternatively, a hybrid approach could be used where there are holistic criteria. (Sadler 2009). Students tend to prefer analytical rubrics which allow students to evaluate each element of their assessment, however lecturers prefer holistic grading (Biggs and Tang 2011; Prins et al 2017)

There also can be a number of stakeholders involved in the design of rubrics, they can be designed with students (bottom up) or by teaching staff (top down) (Prins et al, 2017). Traditionally rubrics have been designed by teachers for students however as the dialog has shifted towards assessment AS learning students are being advocated as partners and co-creators in the assessment process. Relationships and dialogical frameworks are key (Christie et al., 2015; Brookhart and Chen, 2015; Reddy and Andrade, 2010). However institutional policy may inhibit this particularly if students need to be provided with grading criteria at the start of the semester. Also, often students who perceive their identity as passive in the teaching and learning environment may struggle with playing a more active part.

When designing rubrics teachers and students also need to be aware of who they are being designed for and how they will be used. User designed rubrics are focused on supporting both the student to plan and self-assess their work and the assessor to evaluate, mark and provide feedback to students. Assessor designed rubrics are focused on grading the students work with a view to particular quality standards. They are not intended for student use. (Weigle 2002)

Dawson 2017 outlined 14 design elements for rubrics considering level of specificity, scoring and evaluative criteria, language, dissemination and use.

Despite the increasing emergence of research in the adoption and use of rubrics over the past 20 years there are still some limitations of their use (Dawson, 2017). In particular how students interpret and translate rubrics and how they are situated and influenced by social practice and institutional values rather than the transfer of learning to a wider society. Värlander (2008) Which is particularly important

for preparing teachers as researchers, furthermore the narrow view as rubrics as measurement and evaluation tools rather than dialogical frameworks limits their application.

How Students Use Rubrics

A significant amount of research has been conducted on how rubrics impact student learning. Rubrics can improve academic performance and learning (Andrade and Du, 2005; Panadero and Romero 2014)

Also, student soften use rubrics in a strategic manner to enhance their learning for example in planning their assignments, for self-assessment, understanding and satisfying teachers requirements and interpreting teacher feedback. (Bell et al., 2013; Menendez-Varela and Gregori-Giralt, 2016). There is a debate on whether rubrics can increase learner agency and self-regulation. (Alonso-Tapia, and Huertas, 2017) This may depend on the type of student, their experiences and the role of the teacher in integrating assessment and rubrics into the learning process.

Furthermore, rubrics can impact the affective dimensions of learning with students highlighting that it increases confidence and reduces anxiety associated with assessment (Jonsson, 2014; Andrade and Du, 2005). Overall students receive rubrics positively, perceiving them as enabling assessments to be accessible, fair and supporting the quality of their work (Leader, and Clinton, M.S., 2018; Andrade and Du 2005). Although, these experiences often depend on the level of experience students have in engaging in higher education (HE) processes with more experienced learners using rubrics more effectively. Jonsson and Svingby 2007; Panadero and Jonsson 2013 . This may be because often students find the academic language used in rubrics difficult to understand and decipher (Menéndez-Varela and Gregori-Giralt, 2016). As learners engage in HE they become familiar with this language and so use rubrics more effectively.

The research has identified that the adoption of rubrics is not without their limitations. In some cases, particularly for creative disciplines such as the arts rubrics can suppress creative flair (Chapman and Inman, 2009). Furthermore, over prescriptive rubrics are often used in a mechanical way by students who adopt a tick box approach to self-assessing their work (Torrance, 2007; Nicol and Macfarlane-Dick, 2006)

Despite this the success of rubrics is largely dependent on how teachers embed and use them within the teaching and learning environment. The next section will explore how teachers use rubrics in their classrooms.

How Teachers Use Rubrics

The impact of rubrics is highly contextual on how the rubrics are used as and for learning. Many HE educators use rubrics in a very crude manner mainly for the purposes of summative grading and to make the grading process more efficient and consistent. (Bharuthram, 2015) For many it does not form an integral part of their learning and teaching, they feel that the provision of a rubric alone provides sufficient explanation of the assignment expectations for the student. However, rubrics need to be integrated within teaching and learning to support dialogue and enable students to recognize assessments of a high academic quality, self -regulate and become independent learners, particularly considering the constraints learners face regarding academic language. (Panadero and Jonsson 2020)

To aid this, rubrics need to be used for not only summative grading but also formative assessment and feedback. Adopting rubrics in this manner may afford opportunities to enable dialogue between students

and their supervisors and encourage them to develop their identity as partners in assessment (and as researchers) and support relationships between supervisors, students and their peers. Kilgour et al. 2020.

It is important that formative rubrics are developed and used in a flexible dialogical manner. The balance between rigidity and flexibility is key to enable students to incorporate creativity, critical thinking and decision making into their assignment. Through incorporating assessment literacy activities and providing instructional opportunities for students to engage with, the assessment teachers support learners to construct meaning from the assessment and interpret the language used to explain criteria and quality standards (Smith et al 2013)

There are many activities that can support assessment literacy. For example, providing opportunities for learners to engage with the rubric by reviewing samples of assessments and generating feedback using this and discussing how the feedback can be applied to improve the sample; or using the rubric to support student to peer review other learners work and create feedback. Furthermore, when students traditionally engage in peer evaluation they are noted for providing feedback that lacks criticality. Through the use of a rubric students were more likely to provide valuable and critical feedback to support their peers (Pandero et al, 2013)

However, despite these opportunities lecturers still experience challenges when using rubrics particularly for grading. Similar to students some teachers feel rubrics limit creativity and cannot measure soft skills or transferable skills. When using rubrics for analytical grading teachers often find the grade calculate to that of the overall perceived grade and there is a discrepancy between the two, this suggests that some analytical rubrics do not capture or evaluate intangible elements associated with quality academic work. (Bharuthram, 2015; Sadler, 2009). This is particularly evident in research work whereby creativity, problem solving and criticality are key. The next section will explore the role of rubrics in the supervision process.

Rubrics in the Supervision Process

To date the majority of research regarding rubrics has been conducted with undergraduate students, there is a dearth of literature exploring how postgraduate students and their supervisors engage and use rubrics particularly in the context of supporting their dissertation and research theses.

Feedback is a key element of the supervision process. Many students have never engaged in a research project and have no idea of what a 'good 'dissertation looks like (most will have written essays, produced presentations, reports etc.). Through this process students are reliant heavily on academic standards, language and conventions. There is a strong affective dimension associated with conducting research for the first time, with students having a significant amount of anxiety regarding the process. Furthermore, as learners often engage in research alone they find it isolating and unless strong relationships have been developed with other peers and there are opportunities for dialog it can increase apprehension. The learners' experience is heavily reliant on the supervisor relationship, this experience can vary and is not consistent as can be seen in the previous section. Therefore, it is important to provide transparency and support learners to understand the process, their identity and expectations.

Frans et al. 2017 conducted a study with 200 undergraduate research students on how they used and perceived feedback during the supervision process of their theses. They found that students used the rubric similar to other undergraduate students to understand the task requirements, and to evaluate their work. The relationship between the rubric and feedback perception was stronger with students finding it useful for fair and relevant feedback from their teachers and peers.

Within the postgraduate supervision context rubrics have largely been used for summative evaluation of students work, although some supervisors use it for formative purposes. Research regarding formative rubrics illustrated these are often applied a flexible manner with the supervisor reviewing some criteria and omitting others. (Postmes et al 2022). This relates to De Kleijn et al (2015) as 'adaptive research supervision' whereby supervisors nuance their approach based on the student's needs. It is widely recognised that supervisors need to be flexible and adapt their feedback and approach depending on the students' individual context (Chugh, et al, 2021). Rubrics can play a key role in guiding quality of work and providing a framework for dialog and shared understanding which is key to the supervision process. (Kearney 2013). In some cases, supervisors have difficulty in tangibly explaining quality criteria and supervisors often omit specific criteria or make passing comments such as excellent without a rationale for such . (Postmes et al 2022) Key to supporting is meta-dialog which encourages supervisors and students to reflect on the meaning of feedback and how it might impact the students work.

Smith et al, 2013 argued the importance of introducing assessment literacy activities to support meta-dialog about feedback, in which the students and teachers discuss their beliefs and rationale underpinning the feedback so students can bridge the gap between their performance and academic standards. This research was conducted at an undergraduate level however its relevance pertains to the postgraduate supervision process as often students find it difficult to interpret and apply feedback. Therefore, meta-dialogue is key, in this context students need to take ownership of the feedback process questioning the feedback and initiating discussion with the supervisor. However often students do not see themselves in this role, they identify themselves as receivers of feedback rather than joint creators of feedback and feedback as a negotiated dialogical process.

The purpose of this research is to explore

- How students and supervisors interpret and engage with formative rubrics during postgraduate research? And
- How does formative rubrics support the supervisory process during postgraduate research? in particular supporting dialogue and relationship building through the development of goals, discussion of views and feedback, agreeing tasks and nurturing a bond (Bordin, 1993;Dryton 2008)

THEORETICAL FRAMEWORK

As proponents of sociocultural learning theory argue individual cognition proceeds interaction and learning in social contexts. Drawing on Dryton's adapted version of Bordin's (1983) 'Supervisory Working Alliance' Model and we explored the role of the assessment rubric within the supervisory process. This model was developed in the context of counselling to support psychotherapists to assist their clients to incorporate changes. The framework was generalised so it could be adopted in other relational contexts to encourage change for example student-teachers. The framework supports the notion of change which is core to the concept of the teacher as researcher in that it engages teachers in a perpetual quest to understand and improve themselves and their practices (OECD, 2013). Thus, supporting evidence-based changes to practice are central to the identity of teacher as a researcher.

Bordin, 1983 suggested that supervisory work is dependent on the strength of the relationship between supervisor and student and the tasks required to support the change. It involves three core components: goal agreement, task agreement, and bond. Goal agreement involves a shared understanding of the goals

for change; task agreement, a mutual understanding of and confidence in the activities that will achieve these goals; and bond is the emotional attachment between the parties that arises through their work together. This latter affective dimension has been largely overlooked in previous research on the use of assessment rubrics. Dryton (2008) subsequently incorporated a fourth component 'views' which are the beliefs and understandings both the supervisor and student have and how they impact the relationship and the ability to change. The research sought to explore 'How students and staff interpret and engage with rubrics to support the supervisory process during postgraduate research?"

Dryton's extended version of Bordin's (1983) Supervisory Alliance model was adopted as a theoretical framework for the project. The elements of the frameworks were used to inform data collection and analyse the data to determine how rubrics influence the development of goals, forming of views, agreeing tasks and supporting the development of a bond between supervisor and student.

METHODOLOGY

Overview

This research sought to explore learners and faculties experiences of engaging with formative rubrics and how they can be used to support the supervisor-student relationship and the development of teachers as professional researchers in their context. Based on the literature it is clear that faculty supervising research projects often have difficulty in articulating feedback, there is often a lack of consistency in the formative process and rubrics are largely used summative and for grading purposes. (Postmes et al, 2022). This can lead to ambiguity for the student and therefore impact the student-supervisor relationship. How students and staff interpret and engage with formative rubrics during postgraduate research? And How does formative rubrics support the supervisory process during postgraduate research?

The study was conducted with three faculty (the authors of this paper) and three students (authors of this mapper) engaging in a dissertation module in a Master of Education. The purpose of the module is to prepare teachers as researchers in their professional contexts. The dissertation process involves close collaboration between students and their supervisors (faculty members and authors of this paper) over a six-month period. Throughout the process an assessment rubric was used by students to self-assess and generate feedback regarding their own work and with faculty to provide feedback to the students. Both written and oral feedback was provided with students and faculty meeting approximately once per month to discuss the feedback.

This study adopted a Collaborative Auto-Ethnographic (CAE) methodology. Autoethnography seeks to understand the experiences of the self and the social context in which they present in a systematic manner (Ellis et al 2011). CAE allowed the authors to capture both the student and faculty voices and perspectives in a meaningful way (Hernandez et al, 2017) Furthermore it supported the exploration of how learners and faculty engage with rubrics on an ongoing basis throughout the assessment and feedback process. (Adams et al, 2017). In particular the authors felt that identity would play a significant role in how learners and faculty engaged with the rubrics. CAE enabled the authors to explore the role of identity and the relationship between faculty co-authors and students co-authors in how they engaged with rubrics and supported the student - supervisor relationship (Butz and Besio, 2009; Hernandez et al, 2017)

Implementation

Students undertook a supervised research project between November 2021 and June 2022. Students had a variety of supports during the period.

1. Formative and summative rubrics were developed for the dissertation module. To develop these a series of sample rubrics gathered from a variety of third level institutes was presented to an assessment subcommittee of which the faculty coauthors were members, these were formulated into a summative rubric for grading. The faculty coauthors then expanded the summative rubric into formative rubrics to enable them to use these for supervision and formative assessment and feedback. The objective of introducing formative rubrics was to encourage assessment AS learning (National Forum 2016), support meta-dialog between the student and their supervisor (Smith et al, 2019), enable student to continually critique work (Nicol, 2009)The rubrics were presented and discussed with the supervisors and the supervisors met in a community of practice to discuss their experiences of engaging with supervision and how to support students through the process.
2. Students attended regular workshops to support them to complete various sections of their dissertations. These were delivered by the authors; assessment literacy formed a key part of these sessions. The relevant section of the formative rubric was discussed, and students were asked to review sample sections and generate feedback on these samples using the formative rubric. For each section students were asked to complete a peer review of a classmate's work using the formative rubric. This was completed as part of their workshops and provided opportunities for students to revise their drafts before sending to their supervisor. It also facilitated the development of communities of practice within the programme with a view to reducing the isolation experienced during research projects. Furthermore these sessions encouraged shared responsibility and learner agency recognising multiple sources of feedback.
3. Students were allocated a supervisor and sent draft sections to their supervisor ahead of monthly meetings. In these meetings the supervisor and student discussed feedback pertaining to the rubric. The rubric was used to discuss feedback and identify opportunities for enhancement.

Data Collection

The CAE study involved seven student co-authors and three faculty co-authors recording data on how they engaged with the rubric on an ongoing basis. Data was gathered data individually and collaboratively.

1. Rubrics were used to generate written feedback by both the students and faculty.
2. Two collaborative focus groups were held with all ten co-authors (three faculty and seven students) to discuss their overall experiences of the process. This encouraged diverse perspectives and dialogic tensions and enhanced understanding (Sawyer and Norris, 2015)
3. Reflective diaries, all co-authors kept reflective diaries on the after each meeting Bordin's Supervisory Model, 1993.
    a. Goals/practice- what are your goals and motivations for this month – Describe how is the rubric supporting you (or hindering) your progress and the dissertation process as a whole?
    b. Meaning/task agreement: Looking at the rubric in advance of the next section what does it mean for you

c. Identity - What is your role in the supervision process this month, how is it evolving or changing? How do you feel needs to evolve or change for you as a learner?

d. Community/bond – How do you think the supervisor and supervisee relationship could be enhanced? How can the rubric support this?

e. Practice - How do you envisage the rubric supporting the dissertation process going forward in terms of developing as a teacher-researcher?

Data Analysis

Deductive analysis using the theoretical framework outlined in section 4. In particular Drytons, 2008 adapted version of Bordin's (1983) Supervisory Alliance model was used to analyse the written feedback, reflective interviews and focus group data. To minimise bias each researcher analysed their co-researchers and their students' data. Furthermore, inter-coder reliability was adopted with a sample of rubrics and interview data being analysed by two researchers to ensure consistency in results.

FINDINGS

Goal-Oriented Planning and Task Agreement Through Dialogue

The rubric provided clearly defined, well-articulated and challenging criteria designed to guide and scaffold the dissertation journey for both student and supervisor. Rubric use therefore enabled participants to determine goals, define parameters, agree tasks and set destinations. Which aligns to the first component of the supervisory alliance model (Bordin 1993;Dryton 2008)

My goal for the month is to improve my literature review chapter before I begin interviews. I will refer to the rubric throughout this process to ensure that I am on the right track.

(Reflective diary, Student #2).

My goal for this month is to support [student name] in completing her literature review and having a clear idea of where her contribution in terms of the research lies. The rubric has allowed us to identify 'gaps' or areas for improvement and a common understanding of how to support [student name] work going into the next stage of drafting.

(Reflective diary, Supervisor #2).

Goal-setting had three main functions: they provided a general guide to activity, reflection and analysis, enabling students to link their work to rubric criteria; they served as a source of legitimacy, enabling activities to be justified if they contributed to the achievements of identified goals; they were a means of measuring success. The following focus group exchange illustrates these functions:

Student #2: The rubric not only acted as a framework or guide in the writing process, but in our supervisor meetings we would often use it…. as a tool reflection, so we would look at my work through the lens of the rubric and kind of look at what areas were strong, which areas I met, and then which areas

need improvement and how. So, it was also a framework that I would use to look at while I was writing a type of self-assessment so to speak.

Student # 5: I am the type of person where I need to see the checklist. I need to see what's expected of me and then I can kind of work within those guidelines. But I really do need the checklist, the guidelines, the expectations. The rubric did this for me.

(Focus group #2)

Rubric completion by both student and supervisor provided additional direction, scaffolding and feedback to students enabling them to revise drafts based on identified goals. The following excerpt from a completed rubric illustrates such evidence-informed goal-setting and especially illustrates the power of teacher researcher and supervisor dialogue as a driver of professional growth.

Table 1. Rubric from focus group #2

Criterion	Explanation	Student commentary	Supervisor commentary
Relationship to the research question	The review of related research and literature is clearly related to the problem statement as expressed in the research questions and objectives	I am not sure if all my embedded questions are addressed appropriately in this literature review.	I agree. I suggest that you expand/extend the review to incorporate relevant literature/ research associated with RQ3

(Completed rubric excerpt, Supervisor #1, Student #6).

Goal-setting through rubric engagement enabled both students and supervisors to take initiative in improving quality, in affirming and building on what was working well and in identifying areas in need of further development to bring about improvement.

Understanding Views Through Self-Regulated, Transformational Learning

Professional reflection and dialogue, prompted by constructive, rubric-embedded feedback, empowered both students and supervisors to engage in a process of analysis to prioritise for improvement. Key to this was engaging providing students with the ability to clearly define the criteria and challenging them, criteria must also be understood, accepted and implemented in a coherent manner to bring about success. Students revealed that rubrics provided a *focus,* a *benchmark*, a *reference point* to *keep them on track* and motivated them to become self-regulated teacher researchers. This enabled them to understand the views held by the supervisor and the student (D

Student #3: It was like a springboard. I definitely feel like if I was, you know, writing with it there next to me on the desk and constantly referring back and forth to it and learning as a result… It's kind of like a checklist just to check at the beginning, give myself a bit of focus and then check it off again at the end.

Student # 4: The rubric pushed me… Am I reaching the expectation that I expect of myself and how can I improve this? How can I get the best out of my writing for me and I suppose it's kind of trying to push my own right in ability all the time .. to grow as a teacher researcher.

Student #2: You know, my first few drafts, like, I kind of knew they, they might not have necessarily been up to scratch with the rubric. So, it did give me a sort of self-awareness about the quality of the work that I was putting in. And because I was measuring it against the rubric. So, I did find it helpful for that.

Student #1: I'm aware of it and it's kind of like a temperature check for me. So, when I'm writing I'm just kind of looking. Oh yeah, that's the benchmark. That's the standard as opposed to thinking about the assessment, looking at it as an assessment tool. For me it's more.

(Focus group #2)

Supervision meetings provided students and supervisors with opportunities to express their views (Dryton, 2008). In particular their individual opinions, perspectives, experiences, ideas and concerns and to share constructive feedback. These meetings also enabled both students and supervisors to seek further clarification and encouraged students to take a more active role in the supervision process.

I am glad we had an opportunity to explore teacher researcher identity in greater depth today at our meeting. It is apparent to me now that I need to engage more in active listening and open dialogue - to enable [Student name] to uncover the hidden assumptions and misconceptions of her evolving researcher identity.

(Reflective diary, Supervisor #1).

This quote illustrates the central role that dialogue played in a supervisor's ability to construct meaning and to come to a deeper understanding of the dynamic and nuanced process of research supervision. The rubric supported the process of expressing views as it provided a language for the students to ask questions

"....like even in terms of knowing what questions to ask, that echoes with me that like to know that this should be in the methodology chapter or you should do this in the literature review to kind of then be able to ask that. And if that makes sense."

This transformational learning experience, fuelled by a culture of open dialogue and partnership, enabled both students and supervisors alike to take stock and move forward together on the research journey.

Bond Through the Development of Collegial Learning-Focused Relationships

In addition, the rubric supported the development of a bond between the supervisor and student as it illustrated the expectations and framed the feedback that was provided. These relationships were connected, collaborative and increasingly reciprocally developmental. Students and supervisors engaged in processes of meaning creation and sense-making that served to reshape their beliefs, values and sense of 'self'. This transformative experience stimulated personal and professional development, as the following comment reveals:

I think the relationship is very comfortable with my supervisor. I am not afraid to ask for clarification and I appreciate the comments in the learning process as it inspires me to better my research. It also gives me different viewpoints that I have not considered.

(Reflective diary, Student #1).

This exploratory process was negotiated through intense learning conversations, student and supervisor reflections, active listening, trust, warmth and honest collaboration supported by the rubric. These partnerships, involving joint ownership and decision making over both the process and outcome and guided by rubric engagement, stimulated learning and both teacher researcher and supervisor development as the focus group excerpt below illustrates.

Student #3: I mean, it just gave me clarity. I knew what was expected and then after my feedback with [Supervisor name] I knew exactly how to improve. I think it's shaped our conversation and I suppose it maximized our time together because it was very targeted and very focused. There was a context there for the feedback and then I would have gone off and I would have felt armed and motivated to do my second draft on things.

Supervisor #2: But I have to say from a supervisor point of view, I found the reflections after each session very enriching because actually we meet students, we give feedback and we move on but actually having to think about how the meeting went and where it could be improved.. I found really useful …

(Focus group #2).

Collaboration featured as fundamental to constructing meaning and professional growth. Rubric engagement facilitated and fostered a relationship of trust and mutual respect that valued agency, creativity, critical dissent and collaboration. These learning conversations also actively sought experiences of learning with and from students and created safe spaces for risk-taking.

"Supervisor #2: …….it have an insight into the thought process behind the writing, which sometimes then would illuminate for me why that was there or why that was important and perhaps without that I would just be left kind of adding comments not knowing that. That's because of this and that's because of this or that's because I've interpreted this and this particular way or that's why I have used the rubric to inform it in." (Focus group #2).

From a supervisor perspective the rubric enabled the supervisor to understand the students thought process and rationale for decisions which formed a greater understanding of the context of the writing and dissertation process. This allowed the student and supervisor to maximise meeting time and engage in meaningful dialog.

Supporting Teachers in Practice

The rubric also supported teachers in their practice and encouraged them to approach problems or differently. One participant spoke about their experience implementing a new initiative and rather than giving information working with the whole school community to explore how the initiative might work in practice. Furthermore, it empowered teachers in their own practice to research evidence.

"It just shows you that like, no matter what you're doing there, there is information out there. All you have to do is open up the laptop, look for research papers, not necessarily just, you know, a website,

but actually go to the articles, go back to the grassroots of, you know, what the academics or what the researchers have found out." (Focus Group #1)

Furthermore, students saw further potential of the rubric in supporting and mentoring new staff.

"I think as well it the rubric might be helpful, and we do the XXX in our in our school, and I do think a rubric, something like that will be helpful for newly qualified teachers when they have their meetings with their supervisors that they have, I suppose a set of criteria and a rubric, and to know what the they're mentor might be looking for. So, it would be something that I might suggest it in my own school." (Focus Group #2)

It also provided a mindset in which provided teachers with transparency and confidence in their own roles particularly regarding breaking problems down and a language to connect with others in an otherwise often isolated profession. The students felt it is often difficult to connect to peers and the supervision process and rubric gave them the confidence and skill set to reach out to other professionals and work with them on problems facing the profession.

CONCLUSION

This research set out to explore:

- How students and staff interpret and engage with formative rubrics during postgraduate research? And
- How does formative rubrics support the supervisory process during postgraduate research?

Using Drytons supervisory alliance model the researchers investigated how the rubrics supported task agreement, goal setting, bond development and the expression of views. It was found that the rubrics provided transparency in the process, identified expectations and provided a language for students to express and interrogate their work. In particular it provided a meta-dialog to allow students to take ownership of the feedback process questioning the feedback and initiating discussion with the supervisor. (Smith, Worsfold, Davies, Fisher & McPhail,2013). This supported the bonding process between supervisor and student shaping conversations and providing the student with context for the feedback. In addition, the shared reflective diaries which students and supervisors shared provided opportunities to express vulnerability and human elements of the supervision process.

In addition to the elements of Drytons model it was clear that the students used the rubric particularly for planning and self-assessing their work this built their confidence and ensured that they felt their work was to a high academic standard. This aligns to the work of several researchers in particular Andrade and Du 2005; Bell et al., 2013; Jonsson, 2014; Menendez-Varela and Gregori-Giralt, 2016. However, in addition to this the rubric supported them to identify gaps and ask directed questions of the supervisor to allow them maximise the supervisor meetings. In some cases, the students found the language used in the rubric difficult to understand and used lecture notes, sample dissertations and the student handbook to gain a further insight into what the 'terminology' meant again which was a common challenge experienced by student as outlined by Smith et al l,2013. Where this was difficult to decipher they used

the supervisor meetings to probe into the meaning of specific elements of the rubric and how they might reflect this in their work- it formed the basis for the supervisor meetings and conversations.

The rubric adopted was for formative assessment and as a result the supervisors used the rubric to guide conversations, identify gaps and provide structured contextual feedback to the students. This contradicts existing research in which rubrics are largely used for grading and summative purposes. Care was taken to ensure the feedback was reflective of the rubrics. In addition the rubric was discussed in each meeting and so assessment literacy was integrated within the supervisory meetings. The rubric also provided context to the supervisor on the writing process the student engaged in and provided an insight into the rationale behind specific decisions.

Although a small scale study the depth of the qualitative data and narratives provides a meaningful insight into the supervisor and student experiences of adopting rubrics to support the postgraduate supervision process. In particular it is evident that the use of the rubric as a dialogical tool has enriched the process and created a common language for supervisor and students to understand each other's perspectives and enable the development of an open and collegial bond. Furthermore the use of the rubrics in this manner has empowered students to apply their skills as researchers in their teaching contexts and work through open and collaborative approach.

REFERENCES

Adams, T. E., Ellis, C., & Jones, S. H. (2017) Autoethnography. The international encyclopaedia of communication research methods, pp.1-11.

Afdal, H. W., & Spernes, K. (2018). Designing and redesigning research-based teacher education. *Teaching and Teacher Education*, *74*, 215–228. doi:10.1016/j.tate.2018.05.011

Andrade, H., & Du, Y. (2005). Student perspectives on rubric-referenced assessment. *Practical Assessment, Research & Evaluation*, *10*(1), 3.

Bailey, R.A., (2009). Undergraduate students' perceptions the role and utility of written assessment feedback. *Journal of Learning Development in Higher Education*, (1).

Bell, A., Mladenovic, R., & Price, M. (2013). Students' perceptions of the usefulness of marking guides, grade descriptors and annotated exemplars. *Assessment & Evaluation in Higher Education*, *38*(7), 769–788. doi:10.1080/02602938.2012.714738

BERA-RSA. (2014). *The role of research in teacher education. Reviewing the evidence. Interim report of the BERA-RSA inquiry*. London: BERA. https://www.bera.ac.uk/wpcontent/uploads/2014/02/BERA-RSA-Interim-Report.pdf

Bharuthram, S. (2012). Making a case for the teaching of reading across the curriculum in higher education. *South African Journal of Education*, *32*(2), 205–214. doi:10.15700aje.v32n2a557

Biggs, J., & Tang, C. (2011). *Teaching for Quality Learning at University*. Open University Press.

Bordin, E. S. (1983). A working alliance based model of supervision. *The Counseling Psychologist*, *11*(1), 35–41. doi:10.1177/0011000083111007

Boud, D., & Falchikov, N. (2006). Aligning Assessment with Long-Term Learning. *Assessment & Evaluation in Higher Education*, *31*(4), 399–413. doi:10.1080/02602930600679050

Brook, J., Catlin, S., DeLuca, C., Doe, C., Huntly, A., & Searle, M. (2010). Conceptions of doctoral education: The PhD as pathmaking. *Reflective Practice*, *11*(5), 657–668. doi:10.1080/14623943.2010.516981

Brookhart, S. M., & Chen, F. (2015). The quality and effectiveness of descriptive rubrics. *Educational Review*, *67*(3), 343–368. doi:10.1080/00131911.2014.929565

Bui, H. T. (2014). Student–supervisor expectations in the doctoral supervision process for business and management students. *Business and Management Education in HE*, *1*(1), 12–27. doi:10.11120/bmhe.2014.00006

Butz, D., & Besio, K. (2009). Autoethnography. *Geography Compass*, *3*(5), 1660–1674. doi:10.1111/j.1749-8198.2009.00279.x

Chapman, V. G., & Inman, M. D. (2009). A conundrum: Rubrics or creativity/metacognitive development? *Educational Horizons*, *87*(3), 198–202.

Christie, M., Grainger, P. R., Dahlgren, R., Call, K., Heck, D., & Simon, S. E. (2015). Improving the quality of assessment grading tools in master of education courses: A comparative case study in the scholarship of teaching and learning. *The Journal of Scholarship of Teaching and Learning*, *15*(5), 22–35. doi:10.14434/josotl.v15i5.13783

Chugh, R., Macht, S., & Harreveld, B. (2021). Supervisory Feedback to Postgraduate Research Students: A Literature Review. *Assessment & Evaluation in Higher Education*, 1–15. doi:10.1080/02602938.2021.1955241

Cochran-Smith, M., & Lytle, S. L. (1992). Communities for teacher research: Fringe or forefront? *American Journal of Education*, *100*(3), 298–324. doi:10.1086/444019

Cochran-Smith, M., & Lytle, S. L. (1993). *Inside/outside: Teacher research and knowledge*. Teachers College Press.

Cochran-Smith, M., & Lytle, S. L. (1999). Relationships of knowledge and practice: Teacher learning in communities. *Review of Research in Education*, *24*, 249–305.

Cochran-Smith, M., & Lytle, S. L. (2009). *Inquiry as Stance: Practitioner Research for the Next Generation*. Teachers College Press.

Cochran-Smith, M., & Zeichner, K. M. (2005). *Studying teacher education: The report of the AERA panel on research and teacher education*. American Educational Research Association.

Cullen, R., & Harris, M. (2009). Assessing Learner-centredness through Course Syllabi. *Assessment & Evaluation in Higher Education*, *34*(1), 115–125. doi:10.1080/02602930801956018

Darling-Hammond, L. (2016). Research on teaching and teacher education and its influences on policy and practice. *Educational Researcher*, *45*(2), 83–91. doi:10.3102/0013189X16639597

Darling-Hammond, L. (2017). Teacher education around the world: What can we learn from international practice? *European Journal of Teacher Education*, *40*(3), 291–30. doi:10.1080/02619768.2017.1315399

Dawson, P. (2017). Assessment rubrics: Towards clearer and more replicable design, research and practice. *Assessment & Evaluation in Higher Education*, *42*(3), 347–360. doi:10.1080/02602938.2015.1111294

De Kleijn, R. A., Meijer, P. C., Brekelmans, M., & Pilot, A. (2015). Adaptive research supervision: exploring expert thesis supervisors' practical knowledge. *Higher Education Research & Development, 34*(1), pp.117-130.

Doyle, W. (2006). Ecological management and classroom management. In *Handbook of Classroom Management*. Lawrence Erlbaum.

Dwyer, C. A., Millett, C. M., & Payne, D. G. (2006) A Culture of Evidence: Postsecondary Assessment and Learning Outcomes. Recommendations to Policymakers and the Higher Education Community. *Educational Testing Service*. https://files.eric.ed.gov/fulltext/ED500004.pdf

Dysthe, O. (2002). Professors as mediators of academic text cultures: An interview study with advisors and master's degree students in three disciplines in a Norwegian university. *Written Communication*, *19*(4), 493–544. doi:10.1177/074108802238010

Ellis, C., Adams, T. E., & Bochner, A. P. (2011). Autoethnography: An overview. Historical Social Research, 36(4 Elliot, J., (1991). Action research for educational change. OUP: Buckingham

Frischer, J. and Larsson, K., (2000) Laissez-faire in research education—an inquiry into a Swedish doctoral program. *Higher Education Policy, 13*(2), 131-155.

Gatfield, T. (2005). An investigation into PhD supervisory management styles: Development of a dynamic conceptual model and its managerial implications. *Journal of Higher Education Policy and Management*, *27*(3), 311–325. doi:10.1080/13600800500283585

Giroux, H. A. (1988). *Teachers as intellectuals: Toward a critical pedagogy of learning*. Greenwood Publishing Group.

Gurr, G. M. (2001). Negotiating the" Rackety Bridge"—A dynamic model for aligning supervisory style with research student development. *Higher Education Research & Development*, *20*(1), 81–92. doi:10.1080/07924360120043882

Hernandez, K. A. C., Chang, H., & Ngunjiri, F. W. (2017). Collaborative autoethnography as multivocal, relational, and democratic research: Opportunities, challenges, and aspirations. *a/b*. *Auto/Biography Studies: a/B*, *32*(2), 251–254. doi:10.1080/08989575.2017.1288892

Jonsson, A. (2014). Rubrics as a way of providing transparency in assessment. *Assessment & Evaluation in Higher Education*, *39*(7), 840–852. doi:10.1080/02602938.2013.875117

Jonsson, A., & Svingby, G. (2007). The Use of Scoring Rubrics: Reliability, Validity and Educational Consequences. *Educational Research Review*, *2*(2), 130–144. doi:10.1016/j.edurev.2007.05.002

Kearney, S. (2013). Improving Engagement: The Use of 'Authentic Self-and Peer-assessment for Learning' to Enhance the Student Learning Experience. *Assessment & Evaluation in Higher Education*, *38*(7), 875–891. doi:10.1080/02602938.2012.751963

Kilgour, P., Northcote, M., Williams, A., & Kilgour, A. (2020). A Plan for the Co-Construction and Collaborative Use of Rubrics for Student Learning. *Assessment & Evaluation in Higher Education*, *45*(1), 140–153. doi:10.1080/02602938.2019.1614523

Leader, D. C., & Clinton, M. S. (2018). Student perceptions of the effectiveness of rubrics. *Journal of Business and Educational Leadership*, *8*(1), 86–103.

Lee, A. (2012). *Successful research supervision: Advising students doing research*. Routledge. doi:10.4324/9780203816844

Lin, Y. C. (2012). *Online supervision of school counselors: Effects on case conceptualization skills and self-efficacy* [Doctoral dissertation, The University of Iowa].

Lytle, S. L., & Cochran-Smith, M. (1994). Chapter II: Inquiry, Knowledge, and Practice. *Teachers College Record*, *95*(6), 22–51. doi:10.1177/016146819409500602

Mainhard, T., Van Der Rijst, R., Van Tartwijk, J., & Wubbels, T. (2009). A model for the supervisor–doctoral student relationship. *Higher Education*, *58*(3), 359–373. doi:10.100710734-009-9199-8

Menéndez-Varela, J. L., & Gregori-Giralt, E. (2016). The contribution of rubrics to the validity of performance assessment: A study of the conservation–restoration and design undergraduate degrees. *Assessment & Evaluation in Higher Education*, *41*(2), 228–244. doi:10.1080/02602938.2014.998169

Mouton, J. (2001). *How to succeed in your master's & doctoral studies*. Van Schaik.

Moxham, L., Dwyer, T., & Reid-Searl, K. (2013). Articulating expectations for PhD candidature upon commencement: Ensuring supervisor/student 'best fit'. *Journal of Higher Education Policy and Management*, *35*(4), 345–354. doi:10.1080/1360080X.2013.812030

Munthe, E., & Rogne, M. (2015). Research-based teacher education. *Teaching and Teacher Education*, *46*, 17–24. doi:10.1016/j.tate.2014.10.006

National Forum for the Enhancement of Teaching and Learning. (2016). *NF Bulletin: assessment OF, FOR and AS Learning: Students as Partners in Assessment*. Teaching and Learning. https://www.teachingandlearning.ie/publication/students-as-partners/

Nicol, D. (2009). Assessment for learner self-regulation: Enhancing achievement in the first year using learning technologies. *Assessment & Evaluation in Higher Education*, *34*(3), 335–352. doi:10.1080/02602930802255139

Organization for Economic Cooperation and Development (OECD). (2021). *Teaching as a Knowledge Profession: St OECD (2013) Teachers for the 21st Century: Using Evaluation to Improve Teaching*. OECD.

Panadero, E., & Jonsson, A. (2020). A critical review of the arguments against the use of rubrics. *Educational Research Review*, *30*, 100329. doi:10.1016/j.edurev.2020.100329

Panadero, E., & Romero, M. (2014). To rubric or not to rubric? The effects of self-assessment on self-regulation, performance and self-efficacy. *Assessment in Education: Principles, Policy & Practice*, *21*(2), 133–148. doi:10.1080/0969594X.2013.877872

Panadero, E., Romero, M., & Strijbos, J. W. (2013). The impact of a rubric and friendship on peer assessment: Effects on construct validity, performance, and perceptions of fairness and comfort. *Studies in Educational Evaluation*, *39*(4), 195–203. doi:10.1016/j.stueduc.2013.10.005

Parker-Jenkins, M. (2018). Mind the gap: Developing the roles, expectations and boundaries in the doctoral supervisor–supervisee relationship. *Studies in Higher Education*, *43*(1), 57–71. doi:10.1080/03075079.2016.1153622

Pilbeam, C., & Denyer, D. (2009). Lone scholar or community member? The role of student networks in doctoral education in a UK management school. [udying Pedagogical Knowledge across Education Systems. Paris: OECD Publishing.]. *Studies in Higher Education*, *34*(3), 301–318. doi:10.1080/03075070802597077

Postmes, L., Bouwmeester, R., de Kleijn, R., & van der Schaaf, M. (2021). Supervisors' untrained postgraduate rubric use for formative and summative purposes. *Assessment & Evaluation in Higher Education*, 1–14.

Prins, F. J., de Kleijn, R., & Tartwijk, J. V. (2017). Students' use of a rubric for research theses. *Assessment & Evaluation in Higher Education*, *42*(1), 128–150. doi:10.1080/02602938.2015.1085954

Reddy, Y. M., & Andrade, H. (2010). A review of rubric use in higher education. *Assessment & Evaluation in Higher Education*, *35*(4), 435–448. doi:10.1080/02602930902862859

Sadler, D. R. (2009). Indeterminacy in the Use of Preset Criteria for Assessment and Grading. *Assessment & Evaluation in Higher Education*, *34*(2), 159–179. doi:10.1080/02602930801956059

Sadler, R. D. (2014). The futility of attempting to codify academic achievement standards. *Higher Education*, *67*(3), 273–288. doi:10.100710734-013-9649-1

Sarja, A., & Janhonen, S. (2009). Methodological reflections: Supervisory discourses and practice-based learning. *Teaching in Higher Education*, *14*(6), 619–630. doi:10.1080/13562510903315100

Sawyer, R., & Norris, J. (2015). Duoethnography. *International Journal of Qualitative Research*, *8*(1), 1–4. doi:10.1525/irqr.2015.8.1.1

Schmidt, M. & Hansson, E., (2021). "I didn't want to be a troublemaker"–doctoral students' experiences of change in supervisory arrangements. *Studies in Graduate and Postdoctoral Education*.

Smith, C. D., Worsfold, K., Davies, L., Fisher, R., & McPhail, R. (2013). Assessment literacy and student learning: The case for explicitly developing students 'assessment literacy'. *Assessment & Evaluation in Higher Education*, *38*(1), 44–60. doi:10.1080/02602938.2011.598636

Stenhouse, L. (1975). *An introduction to curriculum research and development*. Heinemann.

Stenhouse, L. (1985a). What counts as research. In J. Rudduck & D. Hopkins (Eds.), *Research as a basis for teaching: Readings from the work of Lawrence Stenhouse*. Heinemann Educational Books.

Stenhouse, L. (1985b). The objectives model: some limitations. In J. Rudduck & D. Hopkins (Eds.), *Research as a basis for teaching: Readings from the work of Lawrence Stenhouse*. Heinemann Educational Books.

Sverdlik, A., Hall, N. C., McAlpine, L., & Hubbard, K. (2018). The PhD experience: A review of the factors influencing doctoral students' completion, achievement, and well-being. *International Journal of Doctoral Studies*, *13*, 361–388. doi:10.28945/4113

Tafel, L. S., & Fischer, J. C. (1996). Lives of inquiry: Communities of learning and caring. In G. Burnaford, J. Fischer, & D. Hobson (Eds.), *Teachers doing research: Practical possibilities. Mahwah, N.J. Teaching Council of Ireland. (2017). Initial Teacher Education: Criteria and Guidelines for Programme Providers*. Teaching Council.

Teaching Council of Ireland. (2020). *Céim: Standards in Initial Teacher Education in the Republic of Ireland*. Teaching Council.

Torrance, H. (2007). Assessment as Learning? How the Use of Explicit Learning Objectives, Assessment Criteria and Feedback in Post-Secondary Education and Training can come to Dominate Learning. *Assessment in Education: Principles, Policy & Practice*, *14*(3), 281–294. doi:10.1080/09695940701591867

Värlander, S. (2008). The Role of Students' Emotions in Formal Feedback Situations. *Teaching in Higher Education 13* (2), 145–156.

Wanzare, Z. (2007). The transition process: The early years of being a teacher. In *Handbook of Teacher Education*. Springer., doi:10.1007/1-4020-4773-8_23

Weigle, S. C. (2002). *Assessing writing*. Cambridge University Press. doi:10.1017/CBO9780511732997

Wolcott, W., & Legg, S. M. (1998) *An Overview of Writing Assessment: Theory, Research, and Practice*. National Council of Teachers of English. https://files.eric.ed.gov/fulltext/ED423541.pdf

Zeichner, K. M., & Noffke, S. E. (2001). Practitioner research. In V. Richardson (Ed.), Handbook of research on teaching (4th ed., pp. 298–330).

Zhao, C. M., Golde, C. M., & McCormick, A. C. (2007). More than a signature: How advisor choice and advisor behaviour affect doctoral student satisfaction. *Journal of Further and Higher Education*, *31*(3), 263–281. doi:10.1080/03098770701424983

Chapter 5

The Use of a Rating Scale as a Formative and Shared Assessment Tool in Physical Education

Daniel Bores-García

https://orcid.org/0000-0003-2522-8493

Rey Juan Carlos University, Spain

Raúl A. Barba-Martín

University of León, Spain

Gustavo González-Calvo

University of Valladolid, Spain

David Hortigüela-Alcalá

https://orcid.org/0000-0001-5951-758X

University of Burgos, Spain

ABSTRACT

This chapter presents the analysis of a formative and shared assessment experience in the subject of physical education in secondary education in a high school in Spain. In the assessment process, an assessment scale has been used as an assessment instrument, by means of which students have self-assessed themselves, have been co-assessed by their peers and have received a hetero- assessment from the teacher based on the criteria previously established in the instrument. After the implementation of this experience, a study was carried out on the students' perception of the formative and shared assessment process and the use of the evaluation scale. A discussion group was held, and the teachers' and students' diaries were analyzed. The results show a feeling of motivation and commitment in the students to the task and the group thanks to the feeling of being part of the process through the assessment.

ASSESSMENT, GRADING, AND BODY EXPRESSION

The relationship between assessment and body language as content in the subject of Physical Education (from now on PE) has been a problem for teachers in recent years (Bores-García et al., 2021). On the one hand, the content itself is conflicting from a pedagogical point of view. A careful analysis of this tension reveals three frequent errors that cause this conflict (Lafuente & Hortigüela-Alcalá, 2021): firstly, not developing content that is associated with the collective achievement of the class. Secondly, not generating satisfaction in one's own bodily experiences. Finally, not implementing coherent proposals of a longitudinal nature. This last aspect is one of the main limitations of the projection of these contents since, as they are not applied with a certain logic and continuity, they do not allow learning outcomes associated with the interventions carried out to be obtained (Mattsson & Lundvall, 2015). In this way, it becomes especially complex to generate positive experiences in students since if the PE teacher himself does not give relevance to these contents, students cannot be expected to be motivated towards them (MacLean, 2018). On the other hand, and associated to the difficulty to work on these contents, is knowing how to evaluate them. One of the main problems that exist among PE teachers is the constant identification and homologation of assessment with grading (Álvarez, 2005; Hortigüela-Alcalá et al., 2019a). It should be possible to assess without having to grade everything, and this is a fact that generates many problems in the educational community, especially when working on the contents of PE, which cannot be based solely on surpassing pre-established marks. Assessment, especially in PE, must be synonymous with learning and, for this, the planning must be constructed from the beginning together with the methodological, organizational aspects and the tasks set (Brevik et al., 2017). However, the problem is that when this process is not carried out coherently with the contents to be worked on, the assessment and grading process tends to be hidden from the students (Barba-Martín & Hortigüela-Alcalá., 2022. . In this sense, another problem associated with the confusion between assessment and grading lies in the use of assessment as a power mechanism of the teacher, who justifies being the only one competent to assess because he/she is the one with the knowledge (Fernández-Balboa, 2005, 2007). When this happens, the essence of teaching is lost, as students are not allowed to be an active part of a process of which they are a substantial part and it is difficult for them to develop important aspects of their learning such as autonomy, self-regulation, motivation and awareness of what they have learned((Leenknecht et al., 2020; Ozan & Kincal, 2018). This means that the body does not acquire all the necessary pedagogical dimension within the subject, and that on many occasions the result only responds to physiological criteria and is based strictly on motor performance (López-Pastor et al., 2013). A subject and assessment approach That is detrimental to students and their subsequent physical development outside the classroom (Beltrán-Carillo & Devís-Devís, 2019).

In view of this, a formative and shared approach to assessment are ideal for promotes aspects that relate to and encourage learning such as self-regulation, motivation, responsibility and awareness of learning on the part of the student on the one hand (Hortigüela-Alcalá et al., 2015: Leenknecht et al., 2020; Weldmeskel et al., 2016), and the extrapolation of learning to different contexts on the other (Chng & Lund, 2018). In this line, in the Spanish context in which this work is framed, in the spanish context, in which this work is framed several studies in recent years have shown the benefits associated with these aspects of formative and shared assessment in primary and secondary education (Bores-García et al., 2020; Heras-Bernardino & Herrán-Álvarez, 2019; López-Pastor & Pérez-Pueyo, 2017; Bores-García et al., 2021) and in higher education (Barba-Martín et al., 2020b; Hamodi et al., 2018; Hamodi & Barba-Martín, 2021; Hortigüela-Alcalá et al., 2019b; Romero et al., 2017)., In this way, it is

proven how fundamental it is that in the development of a PE centered on working on the multitude of possible learning of this subject, such as corporal expression, the PE teachers should use assessment as a tool with a clear methodological component, focused on learning awareness and not only on the final grade (Hortigüela-Alcalá et al., 2021). To this end, a clear assessment system and instruments must be developed from the outset (Pegalajar, 2021).

In this sense, the present work tries to contribute a new experience to the literature on assessment in Physical Education and, for this reason, the main objective of this chapter is the narration of a school experience in which formative and shared assessnet was used through the use of an rating scale in a Didactic Unit (DU) of body expression

FORMATIVE AND SHARED ASSESSMENT: AN ESSENTIAL PART OF THE TEACHING-LEARNING PROCESS

But what is formative assessment? and shared assessment? How is it carried out in the teaching-learning process in PE? Formative assessment is an approach focused on the use of this aspect, within the teaching-learning process, to help students achieve quality learning through feedback. (Biggs, 2005; Brown & Glasner, 2003). This assessment model stands out because it is focused on achieving a conscious construction of learning on the part of the students, rather than only on the final result. (López-Pastor et al., 2013). In this sense, it is important to understand that there is no point in grading learning at the end of the process if students has not given the opportunity to learn from their mistakes and to reinforce their successes in order to progress (Brown & Pickford, 2013; Santos Guerra, 2014), which is still too often the case today (Panadero et al., 2019).

On the other hand, shared assessment is a model of assessment that, by understanding this as a formative process, understands as necessary and important the active participation of the students (Barba-Martín & Hortigüela-Alcalá, 2022), based on dialogic, ethical, democratic and transparent relations between the different agents (Santos Guerra, 2003). The learner must be aware of his or her learning and must therefore be involved in the assessment process through different participatory mechanisms such as self-assessment, as the process through which the students or the group is able to evaluate their work at a given moment (Boud, 2013), or peer assessment, where peers give feedback to their peers (Topping, 2009). The possibility for all actors to participate in the assessment of learning processes provides the opportunity for learners to receive feedback from different points of view that help them to become aware of their learning and to be able to direct it. However, it is necessary to establish clear assessment criteria and to create critical scaffolding processes with this type of assessment in the student body, so that all feedback is relevant (Rust, Price & Donovan, 2003). . Under this perspective, the whole assessment process and the instruments must be built, the latter must be clear and provide the students with relevant information from the beginning (Fraile et al., 2017). Some of the most commonly used instruments are rating scales, since they allow students a progressive approach to learning under the construction of evaluation criteria sequenced in different levels of achievement (López-Pastor & Pérez-Pueyo, 2017). Thus, students can know at any moment of the learning process where they are and make decisions based on it. In the field of Physical Education, these evaluation criteria and their levels of achievement must include all areas of student development (motor, cognitive and social) in a related manner. This has provoked a debate in the subject, under the perspective that the importance of the motor aspect is lost. However, this is a false belief supported by the tradition of grading motor contents such as physical condition through tests

and which should be broken in order to provide students with learning processes based on the evaluation and construction of complete motor experiences (Barba-Martín et al., 2020a).

The instruments should be given to the students at the beginning of the requested activity, and their understanding of the criteria and the established achievements should be contrasted with them (Fraile et al., 2017; Pegalajar, 2021). In this way, students can work with the instrument from the beginning and make conscious decisions during the teaching-learning process through self-assessment processes (Nicol & Macfarlane-Dick, 2007). Regarding the feedback received through peer or teacher assessment, the comments received by the students should always be focused on learning, based on the evaluation criteria of the instrument, and with a clear pedagogical intention (Sadler, 2010). Therefore, as can be seen, the use of evaluation instruments with a formative and shared approach is beneficial for student learning; however, despite its advantages, there are still not many experiences and a traditional style of assessment continues to predominate.

A PRACTICAL EXPERIENCE IN PE AT THE SECONDARY EDUCATION STAGE

The work has been developed with 253 students of 1stESO (12-13 years old) in the 2018/2019 academic year and with 182 students of the same grade in the 2020/2021 academic year. The rating scale, inspired by the proposal of Hernando-Garijo, Hortigüela-Alcalá and Pérez-Pueyo (2017), is used during the process of creating and rehearsing a final group performance based on corporal expression and as an instrument of formative evaluation and grading of the DU, using this instrument to perform antriadic assessment (self-assessment, co-assessment and hetero-assessment). The objective of this triangulation of assessments during the DU is for students to contrast the different perceptions of their work in order to regulate their learning in view of the final production. It was decided to use this rating scale because of its high didactic and methodological component. This scale has been extracted from the website of the Attitudes Group (https://grupoactitudes.com/documentos/), which is widely consulted by teachers. This rating scale has been used in many educational contexts, always under the premise of generating awareness and learning in PE students. It should be remembered that the main objective of the research is to assess the experience of PE students when using this assessment instrument, focusing on a triple feedback: self-assessment, co-assessment and heteroassessment. In this case, it has been used in the content of corporal expression. Previous experience in this sense (Pérez-Pueyo et al., 2017;2019) has shown the positive effects of triadic assessment on student involvement and self-regulation of tasks.

The experience narrated is part of the Unit "I use my body to tell stories", belonging to the 1st ESO programme of the subject of PE and timed in the second assessment of the course, between the months of January and February of both years. This DU has been applied in all the groups of 1st ESO of the IES Alameda de Osuna, in the district of Barajas (Madrid, Spain). The unit lasted four weeks, with two 55-minute lessons per week. For its explanation we have divided the DU into phases, according to the assessment work carried out.

Phase 1. Construction of experiences with corporal expression. During the first three sessions, the teacher proposed a series of imitation, representation and expression activities, in which the students worked first individually and then in small groups. The main aim of these first introductory sessions was to arouse the pupils' curiosity about the expressive capacities of the body and to encourage them to explore all the possibilities that could be achieved through movement. Moreover, being the first year of a new educational stage, there was a great deal of heterogeneity in the pupils from the point of view of

their school of origin. This meant that some pupils had deep and positive experiences of content related to body language, while other pupils were confronted almost for the first time with content that not everyone accepts in the same way.

Phase 2. Delivery and explanation of the assessment criteria through the instrument. In lesson 4 the teacher explained to the students the process they were about to start. He detailed the phases of the process and handed out the assessment scale (Table 1), explaining each of the assessable items and answering any doubts that arose in each class. After the explanation, the students were allowed to create their own working groups following two essential criteria: the groups had to be mixed, including at least two boys and two girls, and they could not exceed a maximum number of participants so that there would be five working groups in each class. In the remaining minutes the pupils were allowed to start discussing the first ideas of the creative process.

Phase 3. Creation of the production based on the criteria of the instrument and teacher feedback. During lessons 5, 6 and 7 the pupils created and rehearsed their final performances in groups. Each group had to bring the rating scale to each session, so that they could read the aspects to be evaluated and thus the most important elements to be taken into account.

Table 1. Rating scale

Name and surname of the group members:				
Prerequisites: groups of 4-5 students (mixed). All participants will be involved in the creative process, rehearsals and performance. Three lessons will be dedicated to the creation and rehearsal of the performance.				
Valuable items	**Subvalue**	**Achievement criteria**		**Maximum value**
Storyboard creation	15	The storyboard has good drawings and collects the most important elements of the representation.		15
	10	The storyboard has good drawings and includes some elements of the representation.		
	10	The storyboard has little worked drawings and collects the most important elements of the representation.		
	5	The storyboard has little worked out drawings and picks up some elements of the representation.		
	0	The storyboard has very poor drawings and provides almost no information.		
Duration	10	The performance lasts between 4 and 6 minutes		10
	5	The representation deviates less than 30 seconds from the stipulated time.		
	0	The representation deviates more than 20 seconds from the stipulated time.		
Narrative capacity	20	The story is perfectly understandable		20
	12	The story is understood, although not completely		
	5	There are many elements that are not understood, which makes it difficult to follow the story		
	0	The story is incomprehensible		
Expressive capacity	20	The representation is very rich in expressive body elements		20
	12	The representation has some expressive body elements		
	5	Representation has few interesting expressive resources		
	0	The representation does not have any interesting expressive resources		
Creative capacity	20	The performance is very original and creative		20
	12	The performance has some interesting creative aspects		
	5	The representation has almost no creative elements		
	0	The performance is neither original nor creative		
Flow (coordination)	15	There is coordination and rapport between the members of the group		15
	8	Coordination and rapport can be improved		
	0	The members of the group act in an uncoordinated way, without empathy.		
FINAL SCORE (maximum 100 points)				

Each group worked in an autonomous way and the teacher went around the groups solving doubts, contributing ideas and encouraging those groups that were a bit confused in the creative process.

Phase 4. Peer and heteroassessment process with the instrument. In lesson 8 the students performed the performance and the classmates and the teacher had the opportunity to give feedback based on the

rating scale. At the end of each group's performance, the teacher gave a rating scale to each group to evaluate the group they had performed. He also gave the group a document to evaluate their own performance. Finally, the teacher evaluated the group as well. A few minutes were dedicated to generate a process of student participation with the objective of letting each group know those elements in which they could improve, according to what was expressed in the assessment scale.

Phase 5. Reconstruction of the productions based on the feedbacks received. Lesson 9 was aimed at correcting those elements that could be improved, in order to improve the representation for the following lesson.

Phase 6. New assessment process with the instrument and jump to qualification. Finally, in lesson 10, the final presentation of the performances took place, after each of which the same process took place as in session 8. At the same time, the teacher evaluated each group's performance. A few minutes were devoted to a discussion of the elements assessed, after which the qualitative assessments (achievement criteria) were transferred to the score assigned to each of them, making the final sum out of a total of 100 possible points. In this way, the leap from assessment to marking was made, obtaining a final average mark after equal weighting of the mark obtained in the self-assessment, the co-assessment and the teacher's assessment.

In addition, students completed a semi-structured personal learning diary during some days of the unit, in which they answered some questions proposed by the teacher for each session. Day 1 of the personal diary was the last rehearsal before the first performance, day 2 was the day of the performance, which the pupils themselves called "mock" as it was not the final performance, day 3 was the day of the final performance and day 4 was completed at the end of the didactic unit. The following table shows some of the questions that the students had to answer in each session of the didactic unit.

Table 2. Student´s personal diary questions

Suggested questions for each lesson to complete the diary
Day 1. What motivates you most to do a good performance? Knowing that your classmates and the teacher are going to evaluate you, what does it feel like?
Day 2. Has doing this performance simulation helped you? What things have helped you? What things are you going to change and improve?
Day 3. Describe what you think your level of commitment and work in the group has been during the process of creation and rehearsal. Also give your opinion about the level of commitment and work of your colleagues. What have you learned during the process of creation and rehearsal? How has it been useful for you to have the rating scale during the whole process?
Day 4. How have you experienced having to evaluate your classmates? Did you find it difficult? Did you try to benefit or harm them, or did you remain as objective as possible? How have you experienced having to evaluate your own performance? Did you find it difficult? Do you think the grade given to you by your classmates was fair? Why or why not? Do you think the grade given to you by the teacher was fair? Why or why not?

The PE teacher who carried out the teaching of the didactic unit completed a personal diary with reflections on what happened in each of the sessions with the different groups. Finally, two focus groups were conducted with eight of the participating students.

Table 3. Basic script for the discussion groups

Basic script for the discussion groups
How did you live knowing that you were going to be evaluated in three different ways?
What has been your motivation for trying to do well?
Has the mock performance been helpful in improving performance? In what?
Do you think there have been differences in commitment among the members of your group?
How have you been able to evaluate your peers, evaluate yourself and know that you were being evaluated by the teacher?
Do you consider the score obtained by each of the three ways to be fair?

In order to facilitate the reading of the conclusions obtained from the experience and from the personal diaries and the discussion groups, we will proceed to present the most relevant elements exemplified with some extracts from the teacher´s diary (TD), the student´s personal diaries (SPD) and the focus groups (FG). A qualitative methodology was used to gain in-depth knowledge of the students' evaluations of the use of the scale and their involvement in its evaluation. To this end, the research was carried out in the context of the classroom, so as not to distort the real day-to-day practice of the students, thus connecting with their most personal and social dimension (Halquist & Musanti, 2010). Data were triangulated between the three data collection instruments used. This made it possible, through the use of cross-patterns, to assign each text extract to a category of results, thus obtaining the saturation of ideas (Tortorella et al., 2015). The allocation of data to each of the categories and subcategories was thoroughly reviewed by each of the researchers independently. This process allowed for a coding of the data with each of the instruments used, thus showing linearity and coherence in the analysis (Del Río-Roberts, 2011).

1. Increased motivation in the use of formative and shared assessment

One of the most remarkable elements was motivation. First of all, they were very motivated by the possibility of forming their own groups, even if there were some requirements they had to respect.

One of the things that made them most excited when I explained the working procedure was when I told them that they would have a few minutes to form their own groups, as long as they respected the grouping criteria (TD).

Obtaining a good mark that would have an impact on their final grade in the course was also a motivating factor. The prospect of having part of the responsibility for this grade increased their motivation, as it is usually the teacher who is responsible not only for the assessment but also for the marking of the tests given during the course in the different subjects.

The mark also motivates me, because I want to get a good mark like in the first evaluation (SPD).

Knowing that two thirds of the grade came from what we put on ourselves and the grade our classmates put on us, that was quite motivating (FG).

Finally, the pupils were attracted by the content. Although in each class there is always a group of pupils who relate PE exclusively to sport, they were generally enthusiastic about being able to do a

teaching unit on body language. Even more so as the two previous teaching units had been on handball and physical fitness.

I like being able to show that physical education is more than just sport (SPD).

I was a bit fed up with handball and athletics. I love theatre and that makes me want to do my best for the first time in the course (SPD).

2.   Group feedback and second chances

Students are not used to being able to take an assessment test without consequences for the grade. However, formative assessment is based on the idea of not always having to link assessment with marking. Therefore, students were given the opportunity to evaluate themselves and their peers without the pressure of marking.

Being able to do a performance test before the final was very good for us to know what things we needed to improve and make it perfect in the final assembly (FG).

The students acknowledged that they found this session useful, as they identified shortcomings in their expressive creations that they were able to remedy in the following session.

We realized that the performance was too short and did not meet the minimum time requirement. It's one thing when you do it in rehearsal and another when you do it in front of an audience. When you're nervous, you go faster than planned (SPD).

This possibility allowed them to face the performance of the final session with greater confidence and security, knowing that with the corrections and improvements made, the result would always be better than that of the mock session. In addition, the students who had difficulty in being in front of an observing audience were able to have a first experience in order to face the second one with greater confidence.

Because you've already performed in the simulation, you feel safer and more confident to do it in front of everyone (SPD).

It's helped me to get used to performing in public, because it's not the same to do it only when everyone is looking at you (SPD).

The assessment scale has been an advantage for the pupils, who from the beginning of the project to create the representation were able to know which aspects could be assessed and, therefore, which elements they had to focus on in order to create the representation correctly. The pupils themselves admit to having used the rating scale to detect errors and how to solve them.

The rating scale has been very useful to see where the failures were and what we could do to correct them (FG).

However, the teacher acknowledges that some students did not make correct use of the rating scale, although it has been advantageous for those who have used it correctly.

Some of the students had not read the rating scale carefully, even though we spent almost a whole lesson reading and explaining it. Having to use it to assess themselves and their peers has been good for the next two sessions (TD).

3.   Group commitment

As with any group project, problems arise during its development, the resolution of which is an essential part of the project itself and of interdisciplinary and transdisciplinary learning of great educational value. Many of the groups took the expressive project very seriously, to the point that some of them created Whatsapp groups exclusively to talk about the Project.

...we even have a Whatsapp group to talk to and discuss the failures... (SPD).

The students admit that having common objectives has made them aware of the need to contribute individually to the group. Knowing that assessment and marking were carried out as a group has awakened the need to work as a team, knowing that the consequences of individual actions would have repercussions for the group.

The overall level of involvement has been high. Having shared elements such as a common project, a common evaluation and a common rating has served as a glue for individual and group commitment (TD).

All my classmates have also committed themselves and worked just as hard as I did (SPD).

However, other students have complained about the low level of group commitment shown by some members of the group, who have not contributed to the same extent and have instead hindered the process of creation and rehearsal.

There are people who have had a hard time taking the piss out of it (SPD).

As is always the case, there is a small percentage of students with little motivation for the task (and for everything in general too, almost always) who either blow up the work of others or remain absent, blowing it up anyway. This has occurred above all in groups of students who have not been able to group with whomever they wish, resulting in a group that is too heterogeneous, with students who are more marginalized in the class and, therefore, less motivated (TD).

4.   Lessons learned during the process

Most of the pupils refer to social or relational learning, as the most intense process from an educational point of view has been group work and the resolution of problems that occurred along the way.

It has helped me to know more about my classmates and that some people are not what they seem, because they can be very funny, intelligent and even very good people (SPD).

During the process of creation and rehearsal I have learned to listen (SPD).

This learning is not accidental, but the teacher himself who carried out the didactic unit recognises the intention behind it.

I myself have expressly decided not to give so much importance to expressive content, but to use it as a means of achieving other elements of equal or greater value such as respect for oneself and others, effort, mutual commitment, honesty and transparency (TD).

Working in a team forces students to listen to the opinions of others and not to impose their own. The development of empathy has been a key part of the process.

During the process of creation and rehearsal I have learned to listen (SPD).

5.   Lessons learned during the process

One of the key elements of formative evaluation has been the use of the rating scale, as has already been pointed out several times in this chapter. Its presence throughout the process has allowed the groups to work more efficiently, more focused on the expected results and with greater organisational capacity.

It was quite useful to me because that way we could see everything we had to do and so we were quite focused. We've organized ourselves much better than if we'd had nothing (SPD).

The teacher acknowledges, however, that it is not easy to convey to students the importance of using this assessment tool, especially in this first year of secondary school where some students have already used this type of instrument while for others it is a complete novelty.

Trying to show the students the importance of the rating scale has been difficult for me, I admit. They're not used to using paper and pencil in P.E. Some groups forgot to bring the rating scale to the rehearsals and I got tired of reminding them of its importance. But in the end I think it has been very useful for them (TD)

The lack of previous experience in formative and shared assessment has hindered the process of assimilation of this way of working, because students in the first years of secondary school may misuse shared assessment if they are not used to it.

Evaluating my classmates has been a little difficult, because while I wanted to help them I also wanted to be objective (SPD).

Assessing my own performance was strange to me, because we didn't see what we were doing ourselves (SPD)

However, the rating scale has helped them in this process, especially in the aspect of objectivity or thoroughness in evaluations and feedback.

With the rating scale it has been easier to evaluate my colleagues and myself (SPD).

Following the sheet (rating scale) has helped me stay objective (SPD).

Finally, the leap to marking has also involved some conflicts produced by the inexperience of the students in these assessment processes. The less mature students with greater difficulties of commitment and behaviour did not correctly assess and mark their own work or that of their peers, while the majority of students showed great responsibility and maturity in trying to adjust as much as possible to the assessment scale in order to make judgements and to make the leap to the mark. One of the biggest problems for students has been objectivity. Interestingly, and perhaps only by comparison, students point out that the teacher is objective because he/she has the most knowledge on the subject, whereas they cannot be completely objective.

The teacher's grade did seem fair to me because it was the most objective of all (SPD).

I think I have been fair because the grade we have given the other groups has coincided with the grade those groups have given themselves and with the grade the teacher has given them (SPD).

I recognize that we have scored ourselves a little higher than we really deserved, to help ourselves. The teacher´s grade was the fairest (FG).

The concept of fairness appears in the students' diaries and in the discussion groups very often associated with the concept of objectivity, pointing out that the only fair assessment is that of the teacher or any other assessment that resembles the teacher's assessment.

CONCLUSION FROM THE ABOVE EXPERIENCE AND PROPOSALS FOR FUTURE INITIATIVES

The aim of this paper is to narrate an experience of working on corporal expression in Physical Education through the formative and shared use of a graded scale. For this purpose, the process carried out has been narrated in detail, as well as the main reflections of the teachers and students involved during the process. In this way, a new practical experience is contributed to a still brief literature.

After the body expression teaching unit described above, in which formative and shared assessment was used, the results support the findings of other predecessor research, discussed in the first part of this paper. Students reported feeling more motivated and engaged in the task by being part of the assessment process. They recognise the positive sense of knowing from the beginning of the process the elements that would be observed in the final performance, so that during the sessions they can direct their efforts

more effectively and efficiently. In addition, they comment that they have felt more engaged with their group through the use of formative and shared assessment, feeling more involved in the whole process from the first lesson, not only in the final lesson where the leap to marking is made. The use of this type of assessment has allowed them to understand from the very beginning of the process the effect that individual actions have on the rest of the group at an evaluative level develops the group commitment of all members of each team. In addition, a specific contribution of this experience has been with respect to the use of the instrument as a means of scoring by all the agents involved, it showed small differences between their perceptions, the higher levels of achievement were those coming out of the self-assessment processes, followed by the assessment coming out of the teacher assessment. The mechanism that resulted in the lowest levels of achievement was peer assessment, as the students themselves acknowledge that they are more critical in detecting their peers' mistakes than in detecting their own mistakes and try to be more objective, which they believe makes it easier for them to use the instrument.

The use of formative and shared assessment in general and instruments such as rating scales in particular show their educational potential for increasing the motivation and group commitment of PE students, In this case, to content related to bodily expression. the reported experience, as well as previous research, shows that it is clear that it is advisable to involve students in the assessment processes through, among others, instruments applied using triadic assessment. Although this experience highlights how the use of a formative and shared assessment and the use of an assessment scale require prolonged work in the classroom so that students learn to work with objectivity and awareness of the established criteria, both with their own productions and with those of their classmates.

A proposal for future interventions in which formative and shared assessment is used is the joint creation of the assessment instrument, in such a way that the pupils participate even more in the assessment process by having to make decisions on the achievement criteria and this will serve as a scaffolding with the instrument and the criteria, for which they must previously know the content being learnt and priorities some learning over others in order to grade the degree of achievement of the objectives established for the didactic unit. However, it should be borne in mind that this will take at least two or three sessions prior to the start of the creation and rehearsal, so the teacher must assess whether there is enough time for this. In addition, the process will be different for learners with extensive experience in co-creating assessment instruments than for learners with no experience in this area, because as students experience formative and shared evaluation processes on a constant basis and participate actively in the creation of instruments, these times will be reduced thanks to the accumulated experience.

REFERENCES

Álvarez, J. (2005). *Evaluar para conocer, examinar para excluir*. Morata.

Barba-Martín, R. A., Bores-García, D., Hortigüela-Alcalá, D., & González-Calvo, G. (2020). Evaluación formativa con los estudiantes en prácticas para reducir la brecha teoría-práctica en la formación inicial del profesorado. *Educacion Fisica y Deporte*, *39*(1). doi:10.17533/udea.efyd.v39n1a02

Barba-Martín, R. A., Hortigüela-Alcalá, D., & Pérez-Pueyo, A. (2020). Evaluar en educación física: Análisis de las tensiones existentes y justificación del empleo de la evaluación formativa y compartida. *Educacion Fisica y Deporte*, *39*(1). doi:10.17533/udea.efyd.v39n1a03

Barba-Martín, R. A., & Hotigüela-Alcalá, D. (2022). Si la evaluación es aprendizaje, he de formar parte de la misma. Razones que justifican la implicación del alumnado. *Revista Iberoamericana de Evaluación Educativa*, *15*(1), 9–22. doi:10.15366/riee2022.15.1.001

Beltrán-Carrillo, V. J., & Devís-Devís, J. (2019). El pensamiento del alumnado inactivo sobre sus experiencias negativas en educación física: Los discursos del rendimiento, salutismo y masculinidad hegemónica. *RICYDE. Revista Internacional de Ciencias del Deporte*, *55*(15), 20–34. doi:10.5232/ricyde2019.05502

Biggs, J. (2005). *Calidad del aprendizaje universitario*. Narcea.

Bores-García, D., Hortigüela-Alcalá, D., González-Calvo, G., & Barba-Martín, R. A. (2020). Peer Assessment in Physical Education: A Systematic Review of the Last Five Years. *Sustainability (Basel)*, *12*(21), 9233. doi:10.3390u12219233

Bores-García, D., Hortigüela-Alcalá, D., Hernando-Garijo, A., & González-Calvo, G. (2021). Analysis of student motivation towards body expression through the use of formative and share assessment. *Retos*, *40*(40), 198–208. doi:10.47197/retos.v1i40.83025

Boud, D. (2013). *Enhancing learning through self-assessment*. Routledge. doi:10.4324/9781315041520

Brown, S., & Glasner, A. (2003). *Evaluar en la Universidad. Problemas y nuevos enfoques*. Narcea.

Brown, S., & Pickford, R. (2013). *Assessing Skill and Practice*. Routledge.

Chng, L. S., & Lund, J. (2018). Assessment for Learning in Physical Education: The What, Why and How. *Journal of Physical Education, Recreation & Dance*, *89*(8), 29–34. doi:10.1080/07303084.2018.1503119

Del Rio-Roberts, M. (2011). How I Learned to Conduct Focus Groups. *Qualitative Report*, *16*(1), 312–315.

Fernández-Balboa, J. M. (2005). La autoevaluación como práctica promotora de la democracia y la dignidad. En A. Sicilia, & J. M. Fernández-Balboa (Coords.), La otra cara de la educación física: la educación física desde una perspectiva crítica (pp. 127-158). Inde.

Fernández-Balboa, J. M. (2007). Dignity and democracy in the college classroom: The practice of self-evaluation. In R. A. Goldstein (Ed.), *Useful Theory: Making Critical Education Practical* (pp. 105–128). Peter Lang Publishing.

Fraile, J., Pardo, R., & Panadero, E. (2017). ¿Cómo emplear las rúbricas para implementar una verdadera evaluación formativa? *Revista Complutense de Educación*, *28*(4), 1321–1334. doi:10.5209/RCED.51915

Halquist, D., & Musanti, S. I. (2010). Critical incidents and reflection: Turning points that challenge the researcher and create opportunities for knowing. *International Journal of Qualitative Studies in Education : QSE*, *23*(4), 449–461. doi:10.1080/09518398.2010.492811

Hamodi, C., & Barba-Martín, R. A. (2021). *Evaluación formativa y compartida: nuevas propuestas de desarrollo en Educación Superior*. Dextra.

Hamodi, C., Moreno, J., & Barba-Martín, R. A. (2018). Medios de evaluación y desarrollo de competencias en educación superior en estudiantes de educación física. *Estudios Pedagógicos (Valdivia)*, *44*(2), 241–257. doi:10.4067/S0718-07052018000200241

Heras Bernardino, C., & Herrán Álvarez, I. (2019). La evaluación formativa y compartida desde un enfoque competencial. Aplicación práctica en tareas de aula en Primaria y Secundaria. *Revista Infancia, Educación Y Aprendizaje, 5*(2), 568–575. doi:10.22370/ieya.2019.5.2.1777

Hernando-Garijo, A., Hortigüela-Alcalá, D., & Pérez-Pueyo, Á. (2017). El proceso de evaluación formativa en la realización de un vídeo tutorial de estiramientos en inglés en un centro bilingüe. In V. M. López-Pastor & Á. Pérez-Pueyo (Eds.), *Evaluación Formativa y Compartida en Educación: experiencias de éxito en todas las etapas educativas* (pp. 260–270). Universidad de León.

Hortigüela-Alcalá, D., Abella García, V., & Pérez-Pueyo, Á. (2015). ¿ De qué manera se implica el alumnado en el aprendizaje? Análisis de su percepción en procesos de evaluación formativa. *Revista de Investigación Educacional*, *13*(1), 88–104.

Hortigüela-Alcalá, D., González-Víllora, S., & Hernando-Garijo, A. (2021). Do we really assess learning in Physical Education? Teacher's perceptions at different educational stages. *Retos*, *42*, 643–654. doi:10.47197/retos.v42i0.88686

Hortigüela-Alcalá, D., Palacios, A., & López-Pastor, V. (2019a). The impact of formative and shared or coassessment on the acquisition of transversal competences in higher education. *Assessment & Evaluation in Higher Education*, *44*(6), 933–945. doi:10.1080/02602938.2018.1530341

Hortigüela-Alcalá, D., Pérez-Pueyo, A., & González-Calvo, G. (2019b). Pero… ¿A qué nos referimos realmente con la Evaluación Formativa y Compartida?: Confusiones Habituales y Reflexiones Prácticas. *Revista Iberoamericana de Evaluación Educativa*, *12*(1), 13–27. doi:10.15366/riee2019.12.1.001

Lafuente, J. C., & Hortigüela-Alcalá, D. (2021). La percepción de los futuros maestros respecto a la implantación de contenidos de expresión corporal. *Movimento (Porto Alegre)*, *27*, 1–15. doi:10.22456/1982-8918.111735

Leenknecht, M., Wijnia, L., Köhlen, M., Fryer, L., Rikers, R., & Loyens, S. (2021). Formative assessment as practice: The role of students' motivation. *Assessment & Evaluation in Higher Education*, *46*(2), 236–255. doi:10.1080/02602938.2020.1765228

López-Pastor, V. M., Kirk, D., Lorente-Catalán, E., MacPhail, A., & Macdonald, D. (2013). Alternative assessment in physical education: A review of international literature. *Sport Education and Society*, *18*(1), 57–76. doi:10.1080/13573322.2012.713860

López-Pastor, V. M., & Pérez-Pueyo, Á. (2017). *Evaluación formativa y compartida en educación: experiencias de éxito en todas las etapas educativas*. Universidad de León.

MacLean, J. (2018). Teachers as Agents of Change in Curricular Reform: The Position of Dance Revisited. *Sport Education and Society*, *23*(6), 563–577. doi:10.1080/13573322.2016.1249464

Mattsson, T., & Lundvall, S. (2015). The Position of Dance in Physical Education. *Sport Education and Society*, *20*(7), 855–871. doi:10.1080/13573322.2013.837044

Nicol, D. J., & Macfarlane-Dick, D. (2006). Formative assessment and self-regulated learning: A model and seven principles of good feedback practice. *Studies in Higher Education*, *31*(2), 199–218. doi:10.1080/03075070600572090

Ozan, C., & Kincal, R. (2018). The effects of formative assessment on academic achievement, attitudes toward the lesson, and self-regulation skills. *Educational Sciences: Theory & Practice*, 18.

Panadero, E., Fraile, J., Fernández, J., Castilla-Estévez, D., & Ruíz, M. A. (2019). Spanish university assessment practices: Examination tradition with diversity by faculty. *Assessment & Evaluation in Higher Education*, *44*(3), 379–397. doi:10.1080/02602938.2018.1512553

Pegalajar, M. del C. (2021). La Rúbrica como Instrumento para la Evaluación de Trabajos Fin de Grado. REICE. *Revista Electrónica Iberoamericana sobre Calidad, Eficacia y Cambio en Educación*, *19*(3), 67–81. doi:10.15366/reice2021.19.3.005

Pérez-Pueyo, A., Gutiérrez-García, C., Hortigüela-Alcalá, D., & Hernando-Garijo, A. (2017). Video-diario de evidencias de aprendizaje. Infancia. *Educación y Aprendizaje*, *3*(2), 127–132. doi:10.22370/ieya.2017.3.2.711

Pérez-Pueyo, A., Hortigüela-Alcalá, D., & Gutiérrez-García, C. (2019). La exposición oral con pecha-kucha desde la evaluación formativa. En A. Ramírez y M. P. Gutiérrez (Coords.), La evaluación educativa, entre la emoción y la razón (pp. 104-123). Universidad de Córdoba.

Romero, M., Castejón, F., López, V., & Fraile, A. (2017). Formative assessment, communicative competencies and ICT in teachers training. *Comunicar*, *52*, 1–12.

Rust, C., Price, M., & O'Donovan, B. (2003). Improving Students' Learning by Developing Their Understanding of Assessment Criteria and Processes. *Assessment & Evaluation in Higher Education*, *28*(2), 147–164. doi:10.1080/02602930301671

Sadler, D. R. (2010). Beyond feedback: Developing student capability in complex appraisal. *Assessment & Evaluation in Higher Education*, *35*(5), 535–550. doi:10.1080/02602930903541015

Santos Guerra, M. A. (2003). Dime cómo evalúas y te diré qué tipo de profesional y de persona eres. *Revista Enfoques Educacionales*, *5*(1), 1–15. doi:10.5354/0717-3229.2003.47513

Santos Guerra, M. A. (2014). *La evaluación como aprendizaje. Cuando la flecha impacta en la diana*. Narcea.

Topping, K. J. (2009). Peer assessment. *Theory into Practice*, *48*(1), 20–27. doi:10.1080/00405840802577569

Tortorella, G., Viana, S., & Fettermann, D. (2015). Learning cycles and focus groups: A complementary approach to the A3 thinking methodology. *The Learning Organization*, *22*(4), 229–240. doi:10.1108/TLO-02-2015-0008

Weldmeskel, F. M., & Michael, D. J. (2016). The impact of formative assessment on self-regulating learning in university classrooms. *Tuning Journal for Higher Education*, *4*(1), 99–118. doi:10.18543/tjhe-4(1)-2016pp99-118

KEY TERMS AND DEFINITIONS

Assessment Criteria: They determine the degree of acquisition of key competencies and the achievement of the objectives of each teaching and educational stage.

Assessment Instrument: Documents that are linked to an assessment activity, to establish in them the aspects to be assessed, indicating the levels of achievement of each one.

Assessment Scale: Instrument in which different aspects to be assessed are shown and, in addition, each aspect is subdivided into as many aspects to be assessed, each aspect is subdivided into as many decreasing levels of achievement as it is possible to identify in the production to be assessed.

Formative Assessment: Any assessment process whose main purpose is to improve the teaching-learning processes that take place.

Grading: The process in which a numerical grade or grade is assigned to a student's learning or final production.

Shared Assessment: Assessment process that truly and continuously includes the students, understanding them as the active protagonists of their learning.

Triadic Assessment: It is the joint process of self- assessment, co- assessment and hetero-assessment with the purpose of contrasting perceptions and giving more information in the assessment processes.

Chapter 6

Effective Use of Rubrics in Student Evaluation:
Best Practice E–Portfolios

Elena Ramona Richiteanu-Nastase

iD https://orcid.org/0000-0003-0105-1697

Bucharest University of Economic Studies, Romania

Alexandru Robert Mihaila

Bucharest University of Economics Studies, Romania

ABSTRACT

This chapter clarifies first a number of concepts such as evaluation, traditional evaluation methods, alternative evaluation methods, process-centered evaluation, and evaluation of student progress and portfolios as an alternative evaluation method. The authors will approach the concept of rubrics as a very useful evaluation tool, highlighting design and exemplification ways for some more commonly used evaluation methods. Their advantages and disadvantages will be analyzed. After taking into consideration the most important issues and controversies, the authors will analyze an example of good practice, namely the use of rubrics in the evaluation of students with the help of e-portfolios, as an alternative method of evaluation. The last part of the chapter is dedicated to discussions and recommendations for using rubrics in evaluation. Aspects regarding the usefulness, but also the limits of the instrument, as well as ways of further development, will be discussed.

INTRODUCTION

One of the students' main concerns is the grade they obtain. The first meeting with students is often marked by their desire to know how they will be assessed and how they can get the maximum or sometimes the minimum grade. This student concern is a problem that teachers must address to help students meet personal goals and expectations, as well as the subject's goals or competencies.

In this context, evaluation methods and tools must be clearly indicated at the beginning of the courses, as they offer a guide and formative support for students. Furthermore, discussing the proposed rubrics for the main evaluation method in conjunction with the presentation of the syllabus offers greater opportunities for structured, clear, and comprehensive learning. The presentation and discussion of rubrics and the choice of an evaluation method focused on the student's learning process (e.g., portfolio, project) offer the definite advantage of continuous, formative feedback capable of leading the student to in-depth learning.

This chapter proposes the following:

The background clarifies a number of concepts such as evaluation, traditional evaluation methods, alternative evaluation methods, process-centered evaluation, evaluation of student progress, and portfolios as alternative evaluation methods.

The Student Evaluation Using Rubrics contains two parts*: Rubric Design and Implementation* which approaches the concept of rubrics as a very useful evaluation tool, highlighting design and exemplification ways for some more commonly used evaluation methods. Their advantages and disadvantages are also analyzed and *Using Rubrics with e-Portfolios to Enhance Learning* in which an example of good practice, the use of rubrics in evaluating students with the help of e-portfolios as an alternative evaluation method is analyzed. Step by step, the authors highlight the way to build and implement rubrics for computer-assisted training (a compulsory subject in the students' curriculum for students preparing for a teaching career).

The last part of the paper is dedicated to the *discussions and recommendations* as a result of using rubrics in evaluations. Aspects regarding the usefulness and the limits of the instrument, as well as ways of further development, are discussed.

BACKGROUND

Evaluation is the "activity by which information is collected, processed, and interpreted, regarding the condition and operation of a system and the results obtained, the activity that leads to its evaluation based on some criteria and influences the evolution of the system" (Radu, 2000, p. 18). Similarly, evaluation is regarded as the systematic process of collecting, analyzing, and interpreting information to assess the worth, value, or effectiveness of a program, project, policy, product, or process. It involves making judgments about the quality, relevance, efficiency, and impact of the subject being evaluated. The goal of evaluation is to provide evidence-based insights and recommendations for decision-making, improvement, and accountability (Patton, M. Q., 2008).

Evaluation is the process of determining the extent to which intended outcomes are achieved and the factors contributing to the results. It involves gathering data, assessing progress, and making informed judgments about the merit, worth, and significance of a program or intervention. Evaluation helps stakeholders understand the effectiveness, efficiency, relevance, and sustainability of the initiative and informs future planning and decision-making. (Rossi, P. H., Lipsey, M. W., & Freeman, H. E., 2004).

Numerous terms are used to describe different types and approaches to learning evaluation. Different authors (e.g., Scriven, 1991; Manolescu, 2005; McAlpine, 2002; Radu, 2000; Stoica, 2003) analyze the following: formative (it is a continuous process) versus summative evaluation (at the end of a training sequence), informal (performed in a familiar context, without a stake, is a current assessment) versus formal (assessment context is formal, certifying, or external evaluation), and process (focuses on the learning process) versus product (aims at evaluating the products of the activity).

To measure school performance, the teacher has a wide range of **evaluation methods**. The literature (Scriven, 1991, Manolescu, 2005; Radu, 2000; Stoica, 2003; Rossi, P. H., Lipsey, M. W., & Freeman, H. E., 2004) highlights the existence of some **traditional evaluation methods** (e.g., written, oral, or practical) that focus on the result, the product of learning, and **alternative evaluation methods** focused on the learning process and progress (e.g., case study, observation, project, portfolio, essay). Certainly, each of these methods can be described and analyzed extensively, but for the purpose of the chapter, they will be mentioned briefly.

The *oral examination* is one of the most widely used and can be applied individually or in groups of students. The main advantage of this method is the possibility of teacher-student dialogue, in which the teacher can find out what the student knows in addition to how the student thinks, expresses themselves, and copes with problematic situations different from those encountered during instruction. During the oral assessment, the teacher can ask the student to give reasons for their answer to a particular question, and in return, the teacher can also help with additional questions when the student is at a standstill. However, the method also has some disadvantages; it is time-consuming, and it is difficult to select relatively equivalent questions and topics.

The use of the *written assessment* method ensures uniformity of subjects (particularly in length and difficulty) for the students being assessed, as well as the possibility of examining a larger number of students in the same unit of time. This method is to the advantage of emotional students, who express themselves better in writing and give the teacher a better chance of being objective in the assessment by rereading the work done. In general, the written assessment method does not offer the same possibilities for investigating students' performance (e.g., skills, abilities, capacities, competencies) as oral assessment. This *practical assessment method* allows teachers to see at what level students have formed and developed certain practical skills.

In addition to these traditional assessment methods, there are others that are based on teaching methods. They are called complementary or **alternative assessment methods**. *Reports* are used as a basis for discussion on a given topic and are intended to contribute to the formation or development of students' independent work skills; it is also a possible test of the degree to which students have mastered a particular segment of the syllabus, such as a theme or a more complex issue within a theme.

Projects are also a teaching and assessment method. Projects can be used for learning more complex topics (multidisciplinary, interdisciplinary, and transdisciplinary) and for developing students' investigative skills. Reports and projects call for a thorough consideration of the most appropriate topics to be covered by students.

The advantages of alternative evaluation methods include that 1) they evaluate the process and the progress of learning, 2) they are used simultaneously with the teaching and learning process, 3) the evaluation of school results covers a longer period of time (usually one semester), and 4) learning outcomes that can be measured exceed the sphere of knowledge (as in traditional methods) and measure abilities, skills, and even attitudinal value aspects.

These advantages also apply to the portfolio method. The **portfolio** is also considered the "evaluation report" method (Radu, 2000), and it takes into consideration a relatively large time (semester or school year) and covers all student products in terms of progress from one stage to another. Miller and Legg (1993) define portfolio assessment as a specific form of authentic or performance assessment that attempts to measure higher-order thinking skills including the ability to communicate clearly, to make judgments, and to demonstrate specific competencies. The portfolio may represent a true picture of the student to the extent that the learning activities stimulate different aspects of the student's personality.

These include the overall level of training or understanding of the field (concepts, theories), competencies (skills, abilities, attitudes), outstanding results and interests in certain areas, and learning difficulties encountered. Walland and Shaw (2022) draw attention to the distinction between a portfolio as a process and as a product. Winsor and Ellefson, for example, defined a portfolio as a fusion of processes and products. The process of reflection, selection, rationalization, and evaluation, together with [the] product of those processes (1995, pp. 68–69).

Popham (1995) compares the portfolio and the characteristics of a written test and reveals some clear advantages to the portfolio (p. 164). It engages students in evaluation and provides formative feedback, it measures student performance and also the evaluation of their differences, it is a way of collaborative evaluation and self-evaluation, it evaluates complex elements such as acquisitions, effort, and attitude, and it makes a very good connection between teaching, learning, and evaluation.

The portfolio-based evaluation also has the following disadvantages:

- Time-Consuming: The time allocated to evaluation increases because of the multitude and complexity of issues contained;
- Lack of Standardization: Portfolio assessment lacks standardized criteria and scoring rubrics, which can make it challenging to ensure consistency and comparability across different portfolios or evaluators. Also causes increased difficulty in building a grading scale that reflects elements like creativity and originality of a student (this can be changed using rubrics);
- Does not evaluate the basic level of knowledge;
- Difficulty choosing the best themes that contribute to achieving the objectives/competencies of the discipline.

To mitigate these disadvantages, it is important to develop clear guidelines, criteria, and rubrics for portfolio assessment, provide training to evaluators, and ensure a balanced approach that combines portfolio assessment with other evaluation methods to gather a comprehensive view of students' learning outcomes.

Beyond these disadvantages, the use of the portfolio as an evaluation method should track the student's progress and not be used as a summative evaluation method, in which the student just 'checks' the portfolio pieces. When used incorrectly, the portfolio becomes a traditional method of evaluation.

E-portfolios are a collection and reflection of digital files (artifacts) that are shared electronically for the purpose of reflection, comment, and evaluation. Compared to the advantages and disadvantages mentioned for the portfolio-based evaluation method, the e-portfolio shows the main advantage of using ICT resources. The literature (Abrami & Barrett, 2005; Avraamidou & Zembal-Saul, 2006; Walland &Shaw, 2022). Exploring the influence of web-based portfolio development on learning to teach elementary science. AACE Journal, 14(2), 178-205.) mentions many advantages of using e-portfolios, but the main ones are:

- Can be accessed online, anywhere;
- Can be shared with anyone with access to the internet;
- Can be annotated with comments to provide feedback;
- Offer the owner of the portfolio the ability to decide whether comments should be private or shared;
- Offer the ability to store, organize, and reorder contents quickly and easily;

- Provide opportunities to integrate student coursework and multimedia materials and increase their motivation and satisfaction with the outcome;
- Develop their ability to form the basis for collaboration;
- Enhance the potential for the development of information management, self-organization, planning, and presentation skills.

Using an e-portfolio for evaluation reduces stress for students by working in a collaborative, comfortable environment (e.g., they can work in the classroom, on campus, library, or even at home) and at their own pace.

While e-portfolios offer several advantages, they also have some potential disadvantages. Some relate to the portfolio evaluation method, and some to the online environment. Here are a few disadvantages of e-portfolios:

- Technical Challenges: Creating and managing e-portfolios require a certain level of technical proficiency and familiarity with digital tools.
- Access and Infrastructure: Access to technology and reliable internet connectivity are essential for creating and accessing e-portfolios.
- Privacy and Security Concerns: Storing personal information and sensitive data in e-portfolios can raise privacy and security concerns. It is important to ensure that the digital platform or software used for e-portfolios offers adequate security measures to protect the confidentiality of the content.
- Maintenance and Longevity: Ensuring the long-term maintenance and accessibility of e-portfolios can be challenging, particularly if the platform or software used becomes obsolete or if there are changes in digital file formats over time. Regular updates and backups are necessary to preserve the integrity and accessibility of e-portfolio content.
- Time-Consuming: Developing an e-portfolio can be a time-consuming process. Collecting, organizing, and curating the necessary artifacts and reflections, as well as designing and updating the digital portfolio, can require a significant investment of time and effort.
- Standardization Challenges: E-portfolios may lack standardized criteria and assessment frameworks, leading to potential variations in evaluation and interpretation. This can make it difficult to compare e-portfolios across different contexts or institutions and may affect the reliability and validity of assessment outcomes.
- Limited Interactivity: While e-portfolios provide opportunities for multimedia and interactive content, they may not fully replicate the hands-on or interactive nature of certain physical artifacts or performances. This limitation can impact the representation and assessment of specific skills or achievements.
- It's important to note that many of these disadvantages can be mitigated with proper planning, support, and training. Institutions and individuals should carefully consider the trade-offs and address these challenges to effectively leverage the benefits of e-portfolios.

STUDENT EVALUATION USING RUBRICS

Rubric Design and Implementation

Rubrics are evaluation tools that put the criteria for evaluation and a multilevel description of the levels or standards together in a matrix.

A rubric is a learning and assessment tool that articulates expectations for assignments and performance tasks by listing criteria and describing levels of quality (Arter & McTighe, 2001; Stiggins, 2001).

The **evaluation criteria** take into consideration all the characteristics that must be present in the object of evaluation (product of learning or learning process). They indicate the competencies that will be formed, including elements of knowledge, skills, abilities, values, and attitudes that must be demonstrated. More attention should be paid to the format of these descriptions. They should be designed as observable behaviors.

The **descriptors of performance** at different levels provide information about minimum and maximum standards, as well as intermediate standards. Many authors believe that these descriptions should only be made for three levels because larger scales, for example, the five-level Likert scale, provide relatively similar descriptions, which do not mark the presence or absence of behavior. A good alternative is to use a four-level scale (very good/excellent, good, sufficient, and insufficient). It offers nuances in terms of minimum achievement standards (very good/excellent, good, sufficient) and marks failure through a single level (the area of failure to achieve objectives does not require clarification). An example of rubric criteria and scale is presented in Table 1.

Table 1. Rubric example

Evaluation Criteria	Level of performance/scale				
	Performance descriptors Level 1 Very good/excellent	**Performance descriptors Level 2 Good**	**Performance descriptors Level 3 Sufficient**	**Performance descriptors Level 4 Insufficient**	**Score (total/ defalcated)**
Attribute 1					
Attribute 2					
Attribute 3					

According to Stevens and Levi (2013), rubrics contain a task description or a descriptive title of the task students are expected to produce or perform. Second, they have a scale (and scoring) that describes the level of mastery (e.g., exceed expectation, meets expectation, does not meet expectation). Third, components/dimensions students are to attend to in completing the assignment/tasks (e.g., types of skills, knowledge, so on); and in the end, description of the performance quality (performance descriptor) of the components/dimensions at each level of mastery.

Basically, the usefulness of rubrics is that they give students a clear idea of the performance they need to demonstrate and the characteristics of each level. The easiest way to present these rubrics is to accompany the chosen evaluation method and discuss it with students to understand or negotiate some of the requirements that need to be demonstrated.

A rubric can be analytic or holistic. An **analytical rubric** mentions different dimensions of performance and provides levels of performance evaluation for each dimension. A **holistic rubric** describes the general characteristics and attributes of performance and provides a single score.

The use of the analytical rubric facilitates a clearer identification of the progress of each performance criterion, but it takes a long time to build such an instrument. In this sense, it is best used for longer training sequences (e.g., semester evaluation).

The holistic rubric is easier to build but does not provide feedback for learning and development. Analytical rubrics are needed when the teacher gives a topic that includes a high degree of subjectivity (e.g., essay or problem-solving, projects, and portfolios). In addition, the use of rubrics can be very important when students are new to a particular task or type of expression (Bresciani et al., 2004).

Next, the authors present the **design stages of rubrics** and highlight their specificity for the most used evaluation methods.

Regarding the design stages, there are specialists who mention the characteristics that must be presented in the rubrics (Arter & McTighe, 2001; Huba & Freed, 2000) and others who emphasize certain stages and follow each of them (Bresciani et al., 2004; MacKenzie, 2004; Mettler, 2002; Moskal et al., 2000). The common note of the two approaches is the starting point, namely "desired student learning outcomes" (Luft, 1999).

The first stage of the design of rubrics consists of identifying the learning objectives and transforming them into evaluable objectives that must be demonstrated. This aspect requires breaking the competence into observable performances (observable behaviors).

At the university level, it is common practice to use a syllabus. The syllabus also contains the specific and transversal competencies of the respective subject. Using this information, the objectives/competencies should be transformed into the objectives to be assessed. One way to do this is to use Bloom's taxonomy for the cognitive domain (Gogus, 2012). This prioritizes the mental operations required of the student, from the simplest knowledge to the most complex argumentation. It identifies knowledge, understanding, application, analysis, synthesis, and evaluation. For each of these classes, certain action verbs can be used to demonstrate the presence of the respective learning behavior: knowledge (e.g., recognize, identify), understanding (e.g., reformulate, transform, reorganize), application (e.g., use, apply), analysis, (e.g., differentiate, compare, distinguish), synthesis (e.g., create, produce, design), and evaluation (e.g., argue, validate, decide). The evaluation instrument offered to students assesses whether these competencies are present in the curriculum (e.g., skills, knowledge, attitudes).

In the second stage, for each of the observable performances identified in the first stage, the characteristics or attributes that will be evaluated by the respective evaluation instrument must be established.

The third stage establishes the performance levels/scale and the score assigned for each level. There are a number of descriptors that are usually used to denote the levels of mastery (the excellent-to-poor scale, symbols for letter grades).

In the fourth stage, we describe the attributes/characteristics of the performance for each level. This involves a painstaking task of thinking about indicators for each attribute/characteristic.

The fifth stage is testing, piloting, and review.

The advantages of using rubrics are obvious.

The first advantage is that it provides the student with detailed information regarding the criteria and the level of performance that must be demonstrated.

Rubrics also make a diagnosis regarding the level of progress in relation to the evaluation objectives and represent a self-assessment tool along the way, providing clues regarding the student's position in

relation to the maximum standard. Moreover, they are presented as a learning guide; the student can build their work/project according to these criteria and descriptors.

Rubrics also provide constructive feedback on the steps to follow to reach the next level and provide a clear understanding of the grade obtained because the criteria are clear. They also teach the student to become more objective in their own evaluation;

Finally, rubrics provide transparency and confidence in the grading mode and facilitate objective grading. Rubrics guide and support full learning. An argument in this direction is that the student has the steps to follow so that their performance improves. If they find that they are placed at a medium level on certain criteria, they can improve their answer to reach a higher level.

Rubrics encourage reflective practice for both students and teachers. In particular, the act of rubric design highlights one's values and expectations for student learning and the extent to which these expectations are reflected in current classroom practices.

The value of using rubrics is higher for alternative evaluation methods (portfolio, project, case study). Students can also receive these performance criteria and descriptors with the task, allowing the evaluation product (their paper) to be constructed in accordance with these descriptors. Furthermore, the specificity of alternative evaluation methods means they also receive feedback from the teacher so that students have the confidence that their work is heading in the right direction.

The limits of the rubrics take into consideration the degree of pedagogical training of the teacher and the preference for synthetic or analytical rubrics. There is also a desire to standardize these rubrics. For example, some authors suggest using them for oral presentations. Although it seems like a good aspect because this decreases the subjectivity in the evaluation, the use of rubrics must be done in an adapted way (factoring in age, subject, psycho-pedagogical characteristics). For example, for a student in communication and media and a student in education, the rubrics for oral presentation cannot be the same.

Next, the authors present some examples of rubrics for the main evaluation methods. For an oral examination can be used a simple evaluation rubric such as the one presented in Table 2. The evaluation criteria used can be: response content (scientific accuracy, relevance to objectives; this aspect can be more or less described depending on the complexity of the content), content organization, and response presentation.

The teacher will take into consideration whether the content of the answer is correct and complete in relation to the pedagogical objectives and the content taught. For content organization, the logic of the answer, the ability to synthesize, and the use of examples may be taken into consideration. In terms of response presentation, the most important elements are persuasiveness, message adapted to the listener, support of the message accompanied by nonverbal elements (illustrators), and paraverbal elements (e.g., rhythm, volume).

Table 2. Rubric for oral exams

Evaluation Criteria	Level of performance/scale				
	Very good/excellent	**Good**	**Sufficient**	**Insufficient**	**Score 100 points (total/defalcated)**
Response content	● correct and complete ● scientific accuracy ● relevance to objectives	● relevance to objectives ● correct ● - somewhat scientific accuracy	● correct ● somewhat relevant to the objectives	● incorrect answer or scientifically accurate	60 (20 p*3)
Content Organization	● logic of the answer ● synthetic response ● one example	● logic of the answer ● somewhat synthetic response	● logic of the answer ● somewhat synthetic response		20
Response presentation	● persuasive message ● adapted to the listener ● accompanied by nonverbal and paraverbal elements.	● adapted to the listener ● accompanied by nonverbal and paraverbal elements	● somewhat adapted to the listener ● somewhat accompanied by nonverbal and paraverbal elements	● not target the listener ● unaccompanied by nonverbal and paraverbal elements.	20

As in the case of oral evaluation, for the written evaluation, it is necessary to establish some assessment criteria. Based on the criteria identified for the written test (Muster, 1969), such as the substance of the paper (satisfaction of content requirements, 0 to 6 points and presentation of content, 0 - 1 point), the form of the work (style and spelling, 0 to 1.5 points and graphic presentation, 0 to 1.5 points), and the personal factor (exceptional character, impression, originality, sensitivity, 0 to 1 point), the authors propose a rubric with four levels as presented in Table 3.

Table 3. Rubric for written assessment

Evaluation Criteria	Level of performance/scale				
	Level 1 Very good/excellent	**Level 2 Good**	**Level 3 Sufficient**	**Level 4 Insufficient**	**Score (total/defalcated)**
The substance of the paper	● correct and complete ● scientific accuracy ● relevance to objectives	● relevance to objectives ● correct ● somewhat scientifically accurate ●	● correct ● somewhat relevant to the objectives	● incorrect answer or scientifically accurate	60 points
The form of the work	● correct use of style and spelling ● one or more graphic presentations	● correct use of style and spelling ● no graphic ● presentation	● correct spelling ● somewhat correct use of style	● incorrect spelling ● incorrect use of style	20 points
The personal factor	● original content/ form ● sensitivity to content/form	● original or sensitivity to content/ form	● somewhat original or sensitive to content/form	● not original or sensitive	20 points

At the project level, as an alternative assessment method, it is necessary for the teacher to introduce criteria relating to the process (i.e., documentation, use of data, and information in formulating conclusions) and to the product (i.e., the structure of the project, consistency between content and theme, capacity for analysis and synthesis, relevance of conclusions, the novelty of results).

Table 4. Rubric for project

Evaluation Criteria	Level of performance/scale				
	Level 1 Very good/excellent	Level 2 Good	Level 3 Sufficient	Level 4 Insufficient	Score (total/ defalcated)
The substance of the paper	● uses at least three indicated resources ● consistency between content and theme ● correct and complete presentation of data ● scientific accuracy ● incorporates feedback given	● uses at least two indicated resources ● consistency between content and theme ● correct and somewhat complete ● presentation of data ● incorporates feedback given	● uses at least one indicated resource ● consistency between content and theme ● correct presentation of data ● somewhat incorporates feedback given	● doesn't use resources ● insufficient consistency between content and theme	60 points
The form of the work	● correct use of style and spelling ● one or more graphic presentation	● correct use of style and spelling ● no graphic ● presentation.	● correct spelling ● somewhat correct use of style	● incorrect spelling ● incorrect use of style	20 points
The personal factor	● originality of content/form ● -sensitivity of content/form	● originality or sensitivity of content/form	● somewhat originality or sensitivity of content/form	● not original or sensitive	10 points
Response presentation	● persuasive message ● adapted to the listener ● accompanied by nonverbal and paraverbal elements	● adapted to the listener ● accompanied by nonverbal and paraverbal elements	● somewhat ● adapted to the listener ● somewhat accompanied by nonverbal and paraverbal elements	● does not target the listener ● unaccompanied by nonverbal and paraverbal elements.	10 points

Using Rubrics with e-Portfolios to Enhance Learning

This section contains an example of good practice following the design, implementation, and evaluation of e-Portfolios and rubrics for computer-assisted training, a compulsory subject in students' curriculum who are preparing for a teaching career, and a course offered by the Teachers' Training Department from Bucharest University of Economic Studies. A total of 223 third-year students from eleven faculties of Bucharest University of Economic Studies were evaluated using an e-portfolio with rubrics.

As the infrastructure for creating an e-portfolio was chosen, Google Sites Informatics (a free solution offered by Google) was chosen (https://sites.google.com). There are other e-solutions that can be used to create e-portfolios (e.g., Joomla, Moodle, and WordPress), but Google Sites was selected because most of the students have a Google account. Another reason to choose Google Sites is that it has an

easy-to-use interface that does not require a high level of knowledge and experience in using these kinds of e-solutions.

Using e-portfolios forced requires a number of clarifications. The purpose of the e-portfolio should be clarified: what are the competencies required: knowledge, skills, abilities, and attitudes that we want to be formed in relation to the requirements, abilities, and interests of students in the subject. Moreover, the e-portfolio context (e.g., age, specific specialization of students) and content (e.g., designed activities in relation to the established objectives and thematic units of the discipline) should be clarified.

In this context, the e-portfolio is a complete and complex assessment tool that follows the progress in a certain discipline and the student's attitude toward that discipline over a long period of time. The e-portfolio allows for the formation of specific skills (provided in the syllabus) such as: designing a collaborative platform, layout skills training, implementation of collaborative platforms, design and maintenance of a website, and oral presentation of the e-portfolio. The e-portfolio acts like the student's "business card" as a way to show progress (cognitive, attitudinal, and behavioral) in a particular discipline over a longer period of time.

The e-portfolio is more than an assessment tool; it is a means to enhance the student's independent work, acting as a factor in personality development and reserving an active role in learning. Given the diversity of specialization of students and different levels of ITC skills acquired until the third year of college, the theme of the e-portfolio was chosen by students in compliance with minimum standards set by the teacher (e.g., a minimum of five web pages, implementation of at least two gadgets, inserting multimedia objects, using a chat mode and comments, and formatting a page to allow the administrator to load different platform resources). In other words, it was necessary to use a common evaluation instrument for all eleven faculties, so a rubric was used to guide learning and evaluation. Table 5 is the instrument used to evaluate student performance.

These examples of rubrics are models of good practice that can be a starting point for developing a rubric. They should be treated and taken with skepticism as each subject is unique in terms of the objectives/competencies pursued or the age group addressed.

Table 5. Rubrics for e-Portfolios

Evaluation Criteria	Very good/excellent 25 points	Good 20 points	Sufficient 15 points	Insufficient 10 points
Website design • Create a minimum of 5 pages • Insert at least five objects (e.g., images, tables, organization charts, diagrams, graphs, etc.) • Insert at least two gadgets (e.g., Google Calendar, Google Map, Youtube viewer, Web counter)	• Contains more than five pages and objects • Contains more than two gadgets	• Website contains 4-5 pages and inserted objects • Contains at least two gadgets	• Website contains 3-4 pages and inserted objects • Contains at least one gadget	• Website contains less than three pages and inserted objects • Does not contain a gadget
Format page layout • Use at least three different layouts • Create an interactive page that allows uploading documents (course support, projects, applications, movies, etc.) • Correlates website content with the theme • Uses diacritics and lacks grammatical errors	• Uses more than three layouts • Creates an interactive page • Includes relevant and sufficient content for the chosen topic, • Uses diacritics • Lacks grammatical errors	• Uses 2-3 layouts • Creates an interactive page • Includes relevant and sufficient content for the chosen topic, • Uses diacritics • Lacks grammatical errors	• Uses two layouts • Creates an interactive page • Content relevant to the topic • Lacks grammatical errors	• Uses a single layout • Lacks an interactive page • Irrelevant content for the chosen topic • Grammatical errors
Website design implementation and maintenance • Use of a template according to the theme of the website • Links are functional • Site can be accessed using several browsers and, on several resolutions, • Information is well structured • Divides information into sections and pages • Articles are intuitive, complete, and non-redundant	• Template perfectly reflects the theme of the website • All links are functional • Accessible from more than three browsers • Information presented on the website is structured in sections and pages • Articles are intuitive	• Template perfectly reflects the theme of the website • 3/4 of the links are functional • Accessible from three browsers • Information presented on the website is structured in sections and pages and is intuitive	• Template reflects the theme of the website • 1/2 of the links are functional • Accessible from 2 browsers • Information presented on the website is structured in sections and pages and is intuitive	• Template does not reflect the theme of the website • 1/4 of the links are functional • Accessible from a browser • Information presented on the website is not structured in sections, and the pages are incomplete and confusing
(Oral) presentation of the portfolio	• Cursive speech • Exemplary attire • Grammatically correct language • Mastery of basic concepts • Specific technical language is presented with confidence	• Confident and focused on the content transmitted • Answers most questions correctly • Good command of specific technical language	• Shyness at the time of presentation • Difficulty mastering the specific technical language • Series of grammatical and expression errors occur	• Presents content shyly • Impression that information is being read for the first time • Expressions are incoherent, unclear, and besides the subject • Does not master technical language • Many grammatical mistakes
Final grade	Very well (100 - 81 points)	Good (80 - 61 points)	Sufficient (60 – 50 points)	Insufficient (49 – 1 points)

Any evaluation tool, including rubrics, is perfectible. To create an efficient evaluation tool, the proposed rubrics along with the presentation of the syllabus were discussed in the first couple of weeks of the semester. The objective was to offer greater opportunities for structured, clear, and thorough learning.

The use of the e-portfolios in the evaluation activity entitles permanent communication between the teacher and the students; this communication materializes through continuous feedback, in both directions, regarding the content of the taught subject, the degree of understanding of the students' work tasks for evaluation, and possible measures that can be taken regarding the rubrics to ensure the best possible transparency of the evaluation process. Therefore, the use of e-portfolios allows the teacher to model rubrics in a harmonious and efficient way in terms of design, implementation, and evaluation. Thus, following the feedback received from students, a series of changes to the rubrics can be made. For example, the rubric could target the points allocated in the evaluation for certain criteria, reformulate certain evaluation criteria, and even replace or add other criteria.

SOLUTIONS AND RECOMMENDATIONS

As a conclusion, we offer some solutions and recommendations to address the challenges and maximize the benefits of using e-portfolios:

- Training and Support. There are also a number of limitations related to the teaching skills of faculty, who do not know the steps in the correct implementation of a rubric or copy a certain rubric, which is designed for other objectives, subjects, and level of difficulty. To overcome this obstacle, one can follow the steps mentioned in the chapter on the development of a rubric. In addition, the use of a specific rubric designed for the selected assessment method and adapted to the target group and the subject matter can have positive effects. Online courses to improve teaching skills can be found on this topic.
- Clear Guidelines and Rubrics. Focusing on meeting student and subject objectives is a teacher's top priority This can be facilitated by thinking of a rubric, more or less analytically, to guide and support learning.
- Regular Reflection and Revision: Encourage regular reflection and revision of e-portfolios to promote ongoing learning and growth. Prompt individuals to reflect on their strengths, areas for improvement, and goals, and provide opportunities for them to update their portfolios accordingly.

As Walland and Shaw (2022) state, we agree that portfolios can be introduced into teacher education programs internationally to help pre-service teachers (PST) build records of their learning and reflections. The tool allows them to assemble a collection of evidence of their achievements against graduate standards. These e-portfolios may function as digital CVs and support lifelong learning after graduation.

- Technical Infrastructure: Ensuring access to reliable technology and internet connectivity, but also Implementing appropriate privacy and security measures to protect the confidentiality of e-portfolio content. Educate users about privacy settings, data encryption, and best practices for keeping personal information secure.
- Long-Term Maintenance and Accessibility: Establishing plans and protocols for the long-term maintenance and accessibility of e-portfolios and regularly backing up e-portfolio content can ensure easy access even after completion of the program or course.

By implementing these solutions and recommendations, institutions and individuals can enhance the effectiveness and value of e-portfolios as tools for reflection, assessment, and showcasing achievements.

FUTURE RESEARCH DIRECTIONS

The authors consider that the effectiveness of using rubrics for learning is evident but should be further studied, especially with regard to performance scales/levels. The authors believe that the breadth of the scale might limit the level of understanding and guidance of student learning. In short, the wider the scale, the more levels it contains, and the more it provides a range that is too large to guide learning effectively.

CONCLUSION

The authors appreciate that the value of using rubrics is evident, and using rubrics for alternative evaluation methods (e.g., portfolio, project, case study), as shown in the detailed case study, has its merits because students can receive the rubric with the task, allowing the evaluation product (i.e., their paper) to be constructed in accordance with the descriptors. Moreover, receiving feedback from the teacher gives students the confidence that their work is going in the right direction.

REFERENCES

Abrami, P. C., & Barrett, H. (2005). Directions for research and development on electronic portfolios. *Canadian Journal of Learning and Technology*, *31*(3), 1–15. doi:10.21432/T2RK5K

Arter, J. A., & McTighe, J. (2001). *Scoring rubrics in the classroom: Using performance criteria for assessing and improving student performance*. Corwin Press.

Avraamidou, L. & ve Zembal-Saul, C. (2006). Exploring the influence of web-based portfolio development on learning to teach elementary science. *AACE Journal*, *14*(2), 178–205.

Bresciani, M. J., Zelna, C. L., & Anderson, J. A. (2004). Criteria and rubrics. Assessing Student Learning and Development: A Handbook for Practitioners, National Association of Student Personnel Administrators, 29–37.

Gogus, A. (2012). *Bloom's taxonomy of learning objectives. Encyclopedia of the Sciences of Learning*. Springer., doi:10.1007/978-1-4419-1428-6_141

Huba, M. E., & Freed, J. E. (2000). Using rubrics to provide feedback to students. Learner-Centered Assessment on College Campuses. Allyn and Bacon, 151–200.

Luft, J. A. (1999). Rubrics: Design and use in science teacher education. *Journal of Science Teacher Education*, *10*(2), 107–121. doi:10.1023/A:1009471931127

MacKenzie, W. (2004). Constructing a rubric. NETS'S curriculum series: Social studies units for grades 9–12, International Society for Technology in Education, 24–30.

Manolescu, M. (2005). *Evaluarea scolara* [Lexical characteristics of Romanian language]. Meteor press.

McAlpine, M. (2002). *Principles of assessment*. University of Glasgow, Robert Clark Center for Technological Education., http://www.caacentre.ac.uk/dldocs/Bluepaper1.pdf

Mettler, C. A. (2002). Designing scoring rubrics for your classroom. *Understanding Scoring Rubrics: A Guide for Teachers, ERIC Clearinghouse on Assessment and Evaluation*, 72–81. ERIC.

Miller, M. D., & Legg, S. M. (1993). Alternative assessment in a highstakes environment. *Educational Measurement: Issues and Practice*, *12*(2), 9–15. doi:10.1111/j.1745-3992.1993.tb00528.x

Moskal, B. M., & Leydens, J. A. (2000). Scoring rubric development: Validity and reliability. *Practical Assessment, Research & Evaluation*, *7*(10), 1–6. doi:10.7275/q7rm-gg74

Muster, D. (1969). *Metodologia examinării şi notării elevilor* [Lexical characteristics of Romanian language]. E.D.P.

Patton, M. Q. (2008). *Utilization-focused evaluation*. Sage publications.

Popham, J. W. (1995). *Classroom assessment*. Allyn and Bacon.

Radu, I. T. (2008). *Evaluarea în procesul didactic* [Lexical characteristics of Romanian language]. EDP.

Rossi, P. H., Lipsey, M. W., & Freeman, H. E. (2004). *Evaluation: A Systematic Approach* (7th ed.). Sage Publications.

Scriven, M. (1991). *Evaluation Thesaurus* (4th ed.). Sage Publications.

Stevens, D. D., & Levi, A. (2013). *Introduction to rubrics: an assessment tool to save grading time, convey effective feedback, and promote student learning* (2nd ed.). Sterling.

Stiggins, R. J. (2001). *Student-involved classroom assessment* (3rd ed.). Prentice-Hall.

Stoica, A. (2003). *Evaluarea progresului scolar: de la teorie la practica* [Lexical characteristics of Romanian language]. Humanitas Educational.

Walland, E., & Shaw, S. (2022). E-portfolios in teaching, learning and assessment: Tensions in theory and praxis. *Technology, Pedagogy and Education*, *31*(3), 363–379. doi:10.1080/1475939X.2022.2074087

Winsor, P. J., & Ellefson, B. A. (1995). Professional portfolios in teacher education: An exploration of their value and potential. [Taylor & Francis Online], [Google Scholar]. *Teacher Educator*, *31*(1), 68–81. doi:10.1080/08878739509555100

ADDITIONAL READING

Arter, J. A., & McTighe, J. (2001). *Scoring rubrics in the classroom: Using performance criteria for assessing and improving student performance*. Corwin Press.

Berkley, U. C., & the Center for Teaching and Learning. (2022). *Rubrics*. Berkeley. https://teaching.berkeley.edu/resources/assessment-and-evaluation/design-assessment/rubrics

Chan, Z., & Ho, S. (2019). Good and bad practices in rubrics: The perspectives of students and educators. *Assessment & Evaluation in Higher Education*, *44*(4), 533–545. doi:10.1080/02602938.2018.1522528

Ciesielkiewicz, M. (2019). The use of e-portfolios in higher education: From the students' perspective. *Issues in Educational Research*, *29*(3), 649–667.

Dunn, L., Morgan, C., O'Reilly, M., & Parry, S. (2003). *The student assessment handbook: New directions in traditional and online assessment*. Routledge. doi:10.4324/9780203416518

Gallardo, K. (2020). Competency-based assessment and the use of performance-based evaluation rubrics in higher education: Challenges towards the next decade. *Problems of Education in the 21st Century, 78*(1), 61-79.

Jensen, K. (1995). Effective rubric design. *Science Teacher (Normal, Ill.)*, *62*(5), 34.

Walland, E., & Shaw, S. (2022). E-portfolios in teaching, learning and assessment: Tensions in theory and praxis. *Technology, Pedagogy and Education*, *31*(3), 363–379. doi:10.1080/1475939X.2022.2074087

KEY TERMS AND DEFINITIONS

Alternative Evaluation Methods: Methods that are used in a complementary way, focusing on the process and progress of learning, methods borrowed from training.

Descriptors of Performance: At different levels, provide information about minimum and maximum standards, as well as intermediate standards.

e-Portfolio: A collection and reflection of digital files (artifacts) that are shared electronically for the purpose of reflection, comment, and evaluation.

Evaluation Criteria: The characteristics that must be present in the object of evaluation (product of learning or learning process).

Portfolio: A method of instruction and evaluation; a collection of all the products of student activity, covering a longer period of time and providing indicators of student progress in a particular subject.

Rubric: An evaluation tool that puts together in a matrix a set of criteria of evaluation and the multilevel description of evaluation levels/standards.

Traditional Evaluation Methods: Commonly used evaluation methods focused on the evaluation of the product.

Section 2
Developing Rubrics

The process of development with colleagues and with students. How this can be done and guides for collaborative process and evaluation in different contexts?

Chapter 7
The Importance of Student Partnership in Rubric Construction, Discussion, and Evaluation

Allan Stephen Laville
University of Reading, UK

Alexandra J. Hayward
University of Reading, UK

Lindsey Thompson
University of Reading, UK

Victoria Grace-Bland
University of Reading, UK

Yue Yue
University of Reading, UK

ABSTRACT

The chapter explores the importance of utilising student-staff partnerships in the development and evaluation of rubrics. The approach followed is underpinned by the University of Reading principles for student-staff partnerships that centres student voice in the development of Teaching and Learning initiatives. The chapter explores the challenges of engaging students with assessment rubrics and through engaging in listening exercises, the actions taken to remove these barriers to engagement. The chapter provides three case studies that detail practical recommendations to improving student assessment literacy including in-class support for rubrics, additional support outside of the classroom e.g., assessment rubric screencasts and discussion boards, and the importance of co-creation in creating new rubrics. The chapter concludes by detailing the importance of student-staff partnerships in rubric development and evaluation, but also detailing the additional support mechanisms that need to be in place to effectively develop student and staff assessment literacy.

INTRODUCTION

This chapter supports the view that student-staff partnerships are central to advancing rubric construction and evaluation. In order to explore this view, the objectives of the chapter are:

Objective 1: To provide a literature review on student-staff partnerships and to explore the student-staff partnership framework used at UoR

Objective 2: To explore, through student voice, the current difficulties faced in rubric construction, use, and evaluation

Objective 3: To provide tangible case studies on institutional work that detail solutions to the difficulties raised in Objective 2

Objective 4: To provide future research directions and the implications of our institutional work on sector practice.

BACKGROUND

Partnership is widely recognised as a priority area for engaging learners within the HE sector; as evidenced in The Quality Assurance Agency for Higher Education (QAA)'s UK Quality Code, Advice and Guidance: Student Engagement (2018), Office for Students (OfS)'s Student Engagement Strategy (2022), and Advance HE's Student Engagement through Partnership in Higher Education Framework, as part of their Essential Frameworks for Enhancing Student Success (2016). Partnership can have a positive impact in annual student evaluations including the student voice sections of the National Students' Survey (NSS) and UK Engagement Survey (UKES), and is a highly desirable indicator of success in many award providers such as the OfS's Teaching Excellence Framework (TEF) and Student Minds' University Mental Health Charter Framework (2019).

Partnership between students and staff can be defined as a collaborative and developmental process for unlocking student success and academic engagement, whilst enabling institutions to strengthen their commitment to improving the experiences of their students (Healey and Healey, 2019). Partnership in its truest form focuses less on the outcome, and places importance on building a sense of belonging and connection to the academic community through the process of co-creation (Healey et al., 2014). However, Bovill and Bulley (2011) highlight that the outcomes of partnership work, in addition to the benefits afforded by individuals involved, often result in more impactful research or teaching practices. Lastly, the challenge, and opportunity, of partnership working lies within the uncertainty and flipped power dynamics that arise when staff and students shed their traditional roles, and students (who are inherently experts in their knowledge and experiences of being a student), are empowered to be co-developers of their education (Mercer-Mapstone et al., (2019).

At the University of Reading (UoR), there has been an institution-wide focus to strengthen our approach to student voice and partnership, ensuring we move beyond the formal engagement mechanisms for gathering student feedback, and position students as partners within their own educational experience. With a fully established academic representation system and strong relationship with the elected Students' Union Officers, the student engagement activities were largely governed by University quality assurance policies and led by staff. Bovill and Bulley's (2011) ladder of participation distinguishes the need and importance of different types of student engagement, particularly in curricular design, from

student consultation at the lowest level of the scale, to partnership working where students lead change, sitting at the highest level. . At UoR, there was a clear distinction between perceived surface-level consultations, exercises to listen to the 'student voice', and the need for a deeper, more meaningful and active engagement with our students in the teaching and learning space.

Following feedback from staff and students, including official data such as NSS results, which, in 2018, saw the University's results for questions relating to the Student Voice at 7.58% below the sector, with an average of 65.75% of students agreeing with the section statements, a Student Voice and Partnership strategic project was developed to identify and implement a number of key activities to support staff and students to work together in partnership, allowing the student voice to be heard and acted upon, further empowering students to become agents of change. In our 2022 NSS figures, we saw a marked improvement in the area of Student Voice (UoR 70%, sector av. 66.54%), which we attribute to the commitment of students and staff working collaboratively to enhance teaching, learning and assessment practices at the institution, some of which will be discussed in this chapter.

Leading on from Reading's participation in the national HE pilot project commissioned by TSEP (The Student Engagement Partnership) on behalf of HEFCE (The Higher Education Funding Council for England) in 2014-15, and subsequent working groups, the Student Voice and Partnership strategic project was introduced in 2019 and aimed to: Define, manage, oversee and improve Reading's approach to student engagement, voice, and partnership both within and beyond the classroom; To support staff to increase student engagement through best practice student voice, partnership, and communication approaches; and to monitor and review student voice and partnership mechanisms; and to identify and embed new initiatives to enhance student voice thorough student communications, representation, and feedback. The aspiration for student-staff partnerships at UoR is set out in our institutional *Curriculum Framework*, which outlines that programmes are 'designed and delivered to ensure all students can access and actively engage in their blended learning opportunities through student-centred and inclusive approaches to teaching, learning and assessment. Students work in partnership with staff to shape the programme' (University of Reading, 2021).

The University's foundational work on student-staff partnerships was initiated through the Partnerships in Learning and Teaching Scheme (PLanT), which was set up, and continues, to foster co-creation projects between students and staff within academic departments (e.g., Moys et al., 2018). Subsequently, the UoR Principles of Partnership were developed in collaboration with our Students' Union, Student Reps, and professional and academic colleagues. The principles underpin our ethos for student-staff collaboration by stating that partnership:

Principle 1: Is based on values of trust and respect
Principle 2: Is empowering and inclusive
Principle 3: Enables the collaborative development of meaningful change
Principle 4: Creates a sense of belonging to the academic community

These principles guide our staff development provision and curriculum enhancement work, through which we encourage programme teams to work in partnership with students to explore creative ways to enhance the design, delivery, and evaluation of teaching, learning and assessment. By embedding the Principles of Partnership into institutional policy, strategy, and reward & recognition schemes, we build strong foundations from which we create connections across our academic communities. The principles support students and colleagues to develop a shared understanding of the benefits, challenges and impact

of co-creation activities, which contributes towards a sense of belonging to their programmes, schools, and institution (Healey et al., 2016).

To support academic colleagues to embed the Principles of Partnership into their practice, we co-created a partnership guide with our Students' Union, Student Reps and Student Engagement Community of Practice. The *Partnership at the University of Reading: A Guide for working in partnership with students* (University of Reading, 2020) unpacks each of the principles (indicating key considerations and behaviours) and offers practical examples for embedding within colleagues' teaching or professional context (Healey et al., 2014).

In 2020, amid the global pandemic and the subsequent shift in approaches to teaching and learning in HE, we invested in two institution-wide partnership initiatives, locating students at the centre of the design, delivery and enhancement of their programmes and wider university experiences:

The Student Partners scheme employs students to act as pedagogical consultants within each academic department. Student Partners work collaboratively with teaching staff on projects to enhance the student experience, focusing on school T&L priorities aligned to the objectives of the School Teaching Enhancement Action Plans. Student Partners engage in a range of T&L projects, which might include researching and presenting ideas for educational change and innovation; providing real-time feedback and problem solving; and co-creating student-facing resources (e.g., video guides on assessment criteria and step-by-step guides for aligning feedback with rubrics).

The Student Panel exists to challenge and improve the University by putting student voice at the centre of institutional-level decisions and changes. The Student Panel is a diverse team of 50 students from all levels of study and subject areas (representing the broader demographic of the UoR student body), which actively contributes to shaping the direction of strategic projects. The Panel is an open, transparent, and inclusive space for the diverse voices and perspectives of its members.

The Student Panel allows the institution to: gain a deeper understanding of the personal experiences of students; work proactively in partnership with students at an institutional level; complement existing student voice mechanisms at the University; collect meaningful data that will be used transparently to inform strategic decisions; and demonstrate how students have had an impact on our work at UoR. The themes explored by the Panel are aligned to internal and external measures (e.g., module evaluations and the UK National Student Survey), and have included themes such as assessment criteria, wellbeing in the curriculum, the future of blended learning, academic year restructure, and closing the feedback loop.

Exacerbated by the COVID-19 pandemic, one of our biggest challenges to embedding successful student-staff partnerships came with the move to online learning and removal of face-to-face interactions, key to building relationships, trust and respect within a partnership. At UoR, the introduction of the two partnership schemes at this time resulted in many students and staff facing their first partnership engagement in an online environment, so it was imperative to have a strong set of principles underpinning our aspirations for partnership, from which students and staff could build creative and productive connections. Ntem et al. (2020) discusses the importance of replicating the same ideals of face-to-face partnerships (i.e. trust, dialogue, openness) and outlines some of the benefits of online partnerships, which include removing some of the social barriers and anxiety faced by individuals who have not previously have engaged in a partnership and may avoid engaging in in-person partnership activities. . We introduced both partnership schemes to run virtually in the first academic year (2020-21), and the flexibility this allowed was found to be immensely popular with both undergraduate and postgraduate students, attracting a total of 338 applications to the new roles. Taking a hybrid approach to partnership activities, and the flexibility this allows, has also seen an improvement in attendance and the inclusion of

students beyond the 'super-engaged' (i.e., 78% of Student Panel members did not already hold voluntary and/or paid roles at the university).

For institution-wide partnership roles, we use clear communications and targeted messages in promotions and recruitment to ensure students know these roles are inclusive of our student population, e.g., ''we are particularly interested to hear from anyone from an underrepresented background – this could include international students, Black, Asian, or ethnic minority students, LGBTQ+ students, commuter students, those who are the first in their family to attend university, students with disabilities, care leavers, or estranged students''. Being transparent in our aspirations for engaging a broad demographic of students through partnership activities has seen a positive impact on the expanding the diversity of our actively engaged students, enabling us to capture the voices that may traditionally go unheard. In 2020, the Student Panel membership, mapped against the demographic data of the wider student population at Reading, enabled us to over-represent in a number of demographic datasets, including 30% of Student Panellists declared a disability (UoR population 15.3%), 30% Non-EU domicile (UoR population 18%), and 46% identified as BAME (UoR population 34%).

To ensure our partnership schemes are empowering and inclusive, we work to remove barriers to participation by providing structured support to our students and staff partners, including recruitment guidance, training and skills workshops, and project evaluation activities. By formally employing students in paid roles, we demonstrate equity between our students and staff, and enable more sustainable opportunities for students (who are unable financially or are at a disadvantage) to volunteer their time beyond their studies (Mercer-Mapstone & Bovill, 2020).

Evaluating the impact of student-staff partnership activities is key to enabling a culture of teaching and learning enhancement with student voice at its centre, with a particular focus on evidencing the impact on individuals involved in the schemes. We build in regular points for Student Partners to reflect on their experiences and skills development, framed by the Principles of Partnership. Over the academic years 2020-22, 100% of students on the Student Panel agreed *they had the opportunity to work in partnership with the university*, whilst 94% of students on the Student Panel agreed *they had the opportunity to help inform next steps and decisions made by the university*. For our students employed under the Student Partners scheme over the same time period, 97% of Student Partners *felt their feedback and views were respected,* whilst 92% agreed they *were contributing towards meaningful change in [their] department.*

Beyond influencing change and building stronger connections between students, colleagues and the university, we have also seen a rise in the number of staff cases for promotion, Fellowship of the HEA and internal and external awards that evidence student-staff partnership and co-creation. Although the institution-wide schemes are in their infancy, momentum has been gaining across academic schools and departments and the impact of such interventions can be increasingly seen in the development of teaching, learning, and assessment practices across the institution. In the following section, we explore how we have adopted student-staff partnership approaches to reduce the barriers of engaging students in assessment rubrics.

THE CHALLENGE OF ENGAGING STUDENTS IN ASSESSMENT RUBRICS

The availability of marking criteria is one of the key questions on assessment and feedback in the annual UK National Student Survey (NSS), an independent survey completed by final year undergraduates (OfS, 2023). The survey seeks students' views on their satisfaction with their programme, informs league

tables and demonstrates institutional accountability (ibid). The survey's section on assessment and feedback is typically the weakest performing area across the sector (Andrews et al., 2018); prompting significant time and effort by institutions to improve these scores (Buckley, 2021). Top-down approaches to introduce initiatives to improve assessment and feedback scores are common across the sector. such as using rubrics for each module, initiatives to improve assessment design and articulation, and efforts to improve consistency marking. The development of rubrics aims to move marking from a subjective to an objective activity, by clearly delineating the expectations of student performance according to grade classification. By embracing the use of rubrics to articulate standards, universities can be seen to improve: accountability, achievement, transparency and moderation (Ecclestone, 2001) not to mention fairness for students (Graham et al., 2022).

Across HE, in most disciplines, rubrics are used to attempt to articulate the standard, qualities and/ or features expected for each of the assessment criteria, typically presented in a table, ranging from poor through to excellent performance. Rubrics not only demonstrate assessors' expectations to students in advance of submitting their work, but can also serve to promote a consistent application of criteria amongst marking teams (Sadler, 2009), which have necessarily evolved as participation in HE has grown (O'Donovan et al., 2004) and the time available to impart assessment requirements and standards has diminished (O'Donovan et al., 2008). The implementation of rubrics and their quality can be considered somewhat patchy (Reddy & Andrade, 2010) yet are deemed to have a positive effect on student performance (Brookhart & Chen, 2014). Our second case study recommends the use of staff training at an institutional level to seek to improve consistency of understanding and applications of rubrics among staff.

At the University of Reading, in response to the QAA Frameworks for Higher Education Qualifications (QAA, 2014), we convened a working group to develop generic assessment criteria for each level of study for undergraduate and postgraduate taught programmes, with the intention that these would be contextualised by each of our Schools. The outputs were collated into rubrics, highlighting expectations for each grade descriptor, and embedded in our Assessment Handbook (University of Reading, 2018).

Exploring Students' Perceptions of Marking Criteria

Following the development of institutional generic marking criteria, we sought to work in partnership with students to gain a better understanding of responses to the NSS question on "*The criteria used in marking have been clear in advance*". We recognised that the plethora of language pertaining to assessment and feedback can appear confusing to students, given the broad range of terminology offered in the literature e.g., standards, schema, criteria, rubrics (Sadler, 2009) and even differences in the activity undertaken by staff when making judgements e.g., grading, marking, or assessing.

50 student panellists (undergraduate and postgraduate) completed a pre-focus questionnaire to explore their understanding of assessment criteria terminology and their experiences of using such criteria. The questionnaire was developed following a review of contemporary literature and the NUS Assessment Benchmarking Tool (2015). Four focus groups were convened to share and explore the findings of the questionnaire in more detail and to identify areas for improvements.

Students' Perceptions of Assessment Rubrics

Our initial concerns regarding the language of assessment criteria appeared unfounded, many students believed the criteria were easy to understand and helped them understand markers' expectations (86%).

We noted however that students used terminology such as assessment criteria/tables, grading criteria, rubrics, and marking criteria interchangeably in their responses. Furthermore, their understanding of assessment terminology varied e.g., some saw assessment criteria, marking criteria, and/or rubrics as having individual purposes. Assessment criteria were seen to apply to the whole module but marking criteria existed for individual tasks and rubrics were to be used to understand feedback. This conflation of terminology suggests students considered that different standards or schema might be being used to assess their learning. Given the language used in the NSS, this confusion could possibly explain the survey's results, if students perceived marking criteria to be something different to assessment criteria, they feasibly believed "the marking criteria were not available in advance".

When prompted for explanations of their responses about ease of use, we identified a more complex picture. Students liked that the information within rubrics was clearly laid out, and did highlight expectations, it delineated the different skills expected, and highlighted differences between grade boundaries. Rubrics therefore provided the transparency advocated by Fletcher et al. (2012) and Jonsson (2014) but not the assessment literacy they proposed rubrics offer, as students found them difficult to use as a tool to identify what to do to boost their grades. Our second case study aims to explore some solutions to this through the use of multi-modal introductions to assessment using screencasts and discussion boards.

The student panel found descriptors within rubrics were considered too indiscriminate, featured ambiguous and repetitive language between grade bands and lacked specific information about the features of a good essay for a specific programme. Whilst students thought rubrics could be used as a checklist or framework to guide them on what to do; many of these students identified areas for improvements. It appeared students are purposefully seeking clues around the expectations of academic staff and could be perceived to be taking a "*mechanistic approach*" to assessment criteria as described by Norton (2004, p700), rather than focusing on the learning to be undertaken or the skills developed. Our third case study finds that students struggle to see that the same skills are being assessed across different assessments. Graham et al., (2022) suggested students' approach to using assessment criteria as a set of instructions is not the purpose for which rubrics are intended, a misconception echoed by Matshedisho (2020) who flagged a gap between staff's use of rubrics to convey conceptual knowledge and students' engagement with rubrics to seek procedural information.

Generic vs. Specific Assessment Criteria and Rubrics

We invited the student panel to express a preference for generic or task-specific assessment criteria / rubrics. The majority favoured task specific, noting that they needed the context specific information to understand what was expected. Where generic rubrics for a type of assessment were used e.g., essays within a department; this was seen as ambiguous, thus not allowing the use of a rubric as a recipe for assessments (Cockett & Jackson, 2018). The potential for generic criteria to highlight development and improvement through a programme was not recognised, as students favoured task-specific information to point them in the right direction for their next assignment (Jonsson, 2014). Upon further reflection, students conceded that generic rubrics could potentially help first years transitioning into HE, giving them an overview of expectations, how to do university work and improve on it. Our third case study explores how generic rubrics can be used to support staff and students with assessment and feedback practices, and crucially examines the role of dialogue in supporting implementation of such rubrics.

Identifying Assessment Criteria and Rubrics

A third of students highlighted dissatisfaction with their ability to find criteria and be supported to understand them. Interestingly the data maps neatly to institutional NSS scores for assessment criteria availability, which align with the sector benchmark (expected score based on student and course characteristics, (OfS, 2023)). Students told us they could typically look on Blackboard for the assessment criteria [the VLE] but for some students, variation between modules caused frustration and confusion (Andrade & Du, 2007). In rare instances criteria were not released until after the work had been returned, thus acting as a "secret scoring sheet" (Dawson, 2017 p347) and the potential existence of a hidden curriculum, whereby a gap exists between what is intended and the lived experience of assessment by the students (Sambell & McDowell, 1998). Jonsson (2014) noted that by providing access to the rubric in advance, students can have confidence in the assessment process, this would serve institutions well in the NSS. The clear articulation of not just the criteria but the marking process is critical to absolve notions of a hidden curriculum. The procedures for marking school and college work in the UK are significantly different to HE e.g., detailed schema is not produced for each assessed element, marks are not norm referenced in line with previous cohorts. It would be unreasonable to suggest students entering HE understand how their work is marked unless we take it upon ourselves to explain it; thus, increasing transparency, trust and confidence in the process, using consistent language and approaches. Students must know how judgements on their work will be made at the start of their learning; securing their trust by eliminating surprise assessment strategies and clearly communicating processes and expectations (Padden & O'Neill, 2021).

The Language of Assessment Criteria and Rubrics

Although students understood the words within the rubrics, they used terms such as "*vague*", "*subjective*" and "*difficult to differentiate*" echoing Cockett and Jackson (2018) and O'Donovan et al., (2001) but here, for a wider range of disciplines. Despite an emerging recognition of the value of rubrics, the use of language in the rubrics was opaque, leaving students wondering what the difference between good and very good might be? The more specific the terminology was, the more useful the rubrics were thought to be. If terminology within rubrics were to remain centred on good, very good, excellent etc. students thought that examples would enable them to understand the differences between the grade categories, taking them from an abstract concept to something tangible to guide them in understanding expectations of markers. It was clear from the responses that without support, the rubric was not particularly helpful as a tool in assessment for learning, which we further investigate in case study 2. It is worth noting however, that by over-emphasising assessment criteria, students will become too concerned with wanting ever increasingly precise definitions and explanations of the criteria (Norton, 2004) thus distracting them from their learning.

Students' Interpretations of Assessment Rubrics

When it came to support in understanding the assessment criteria, it is clear that the mere provision of criteria/a rubric is insufficient. Jonsson (2014) posited that without training in understanding rubrics, there can be little to no development of students' assessment literacy. Instead, active engagement with the rubrics is required e.g., self-assessment activities which demand interaction with the rubric.

Students found it helpful when staff went through the assessment criteria and unpacked the terminology for them, given that "*Rubrics are not entirely self-explanatory. Students need help in understanding rubrics and their use*" (Andrade, 2005, p.29). A mix of practices were identified after the provision of criteria and rubrics. In some instances, students worked alone to explore what the criteria meant, others attended seminars run by staff to go through the criteria whilst other students approached staff individually. Video guidance and/or screen casts were seen as particularly helpful, enabling students to gain an understanding of expectations and to revisit these as often as they wished. The latter clearly offers a more inclusive approach to supporting students which does not rely on them feeling empowered to seek out tutors, be on campus or attend a class on a particular day. Please see Case Study 2 for how we have embedded inclusive approaches to improving students' assessment literacy.

Without in class support, students tended to see assessment criteria as the guide to what to include and the rubrics as something they would only refer to after they have received their feedback. In some cases, they perceived there to be a disconnect between the criteria indicating what to do and the rubrics showing what they achieved without a clear understanding of the difference.

Supporting Students to Engage with Rubrics

Practice varied between programmes and modules in terms of whether a set of standard generic criteria were used, the type of support offered to students (naturally there will be disciplinary variation, and even within a discipline) and in terms of where to find criteria and rubrics. All but one of the students were very keen for time to be spent in class guiding students on how to tackle their assignment, to understand the standard of work required, and to gain insights into markers' expectations. We outlined some mechanisms for support for the students to comment on, ranging from traditional, more passive approaches to partnership approaches which prompt active engagement to promote a better understanding of the criteria (O'Donovan et al., 2008).

Students considered the **use of VLE discussion boards** offer a pragmatic, efficient and inclusive means of collating and responding to student assessment queries, offering a transparent approach to develop assessment literacy. The lack of "*interaction, immediacy of response*" and opportunity for "further explanation" were noted. The value of using a VLE is further explored in our second case study, whereby we encourage students to engage in dialogue about their assessments through an anonymous Blackboard discussion.

Opportunities to use rubrics to mark exemplars familiarises students with the academic standards for an assessment and improves their capacity to make judgements on a piece of work (Carless & Chan, 2016) but does require time and practice (Andrade & Du, 2007). The lack of practice is perhaps evident in the responses, students expressed frustration that their interpretation of the rubric didn't match that of staff, so it felt like a bit of a waste of time. Concerns exist among the cohort that what they would judge as good might not match tutors' expectations of good work. The dialogic benefits of discussing the rubric in class were recognised, which enabled students gain a deeper understanding of the criteria (Graham et al., 2021).

Self and peer-assessment had not been encountered by all the students in the focus groups, those with no experience thought it would lead to a deeper understanding of the marking process, prompting engagement with it, rather than superficially glancing through good and bad examples of work. The themes emerging from the discussion echo that of Andrade and Du (2007), students did not consider they had the skills to self-assess; with some noting that they thought they would be harsh on themselves. With

time, they believed they would improve their self-assessment, at which point the rubric would become critical to developing their revised drafts due to a deeper understanding of the criteria. Peer assessment was less popular as students were reluctant to share their work, suggesting they would not trust other's ability to mark, thus highlighting the importance of such interventions being formative (O'Donovan et al 2008) and students being taught how to carry out peer review (Mulder et al., 2014) to avoid pitfalls around group politics and bias.

Co-creation of assessment criteria and rubrics aims to create a shared understanding and consensus through partnership and dialogue (Lorber et al., 2019), recognising that students are experts in being students at their institution (Cook-Sather et al., 2014). Most students had yet to encounter opportunities to co-create with staff but deemed that it would be useful as the end-result would be student centred and accessible. This chimes with Graham et al., (2021) who explored the value of such partnership when developing new and/or unfamiliar assessments. The partnership unveiled the boundaries of students' understanding of the rubrics and identified areas where further work was required to create something meaningful for staff and students. As with many surveys into partnership (e.g., Martens et al., 2019), students expressed a concern that they would not necessarily have the confidence, knowledge or skills to co-create with staff.

The student panel on assessment criteria and rubrics provided us with some insights into students understanding of the usefulness or otherwise of assessment rubrics in HE. It is evident that whilst rubrics are useful, they cannot be used in isolation. Clear articulation of expectations, support to unpack rubrics, and assessment support sessions which prompt students to actively engage with rubrics are critical. Furthermore, there is much work to be done to develop students' assessment literacy to enable them to understand the nuances of the language within rubrics. Conversely, there is also an onus on staff to ensure rubrics are accessible, inclusive and user friendly to avoid the creation of a hidden curriculum. We explore these concepts and solutions to the issues raised further in the following three case studies.

SOLUTIONS AND RECOMMENDATIONS

Case Study 1: Psychology — Improving Student Engagement With Assessment and Feedback

Having investigated the institutional perspective on rubrics, we explored the challenges faced by the School of Psychology and Clinical Language Sciences. The School is committed to enhancing the student assessment experience and positive student experience could result from making assessment marking criteria accessible early in the semester and articulating the expectations. The clear communication of assessments criteria could be achieved by presenting assessment briefs and well-developed rubrics, which consist of grade descriptors of marking criteria, quality definitions (from excellent to fail) and a scoring strategy (e.g., Arter & Chappuis, 2007). Rubrics of a variety of assessments (essay, report, presentation, etc.) were specified by our teaching team in accordance with the University guidelines on Foundation, Undergraduate and Postgraduate programmes. In the process of rubric design, the student partnership approach to teaching enhancement was implemented in the working agenda. For instance, student partners (one student research assistant and five student participants) facilitated focus groups that investigated students' use of rubrics and the feedback they received. The teaching team then incorporated student voice that reflected on the experiences of navigating rubrics, developing assessment literacy,

and engaging with feedback. This case study aims to discuss the role of rubrics in the learning journey based on the findings from our recent student partnership projects.

Researchers (e.g., Andrade, 2005; Giacumo & Savenye, 2020) have argued that rubrics can support teaching and learning and serve instructional purposes. Students who consult rubrics as a guidance tool achieve higher grades and deeper learning (e.g., Brookhart & Chen, 2015). Consistent with published studies on rubrics and learning (e.g., Bolton, 2006; Jonsson, 2014), our projects found that students appreciated the transparency of marking schemes and valued rubrics that clarified the requirements for their work. Early access to rubrics enabled students to identify critical issues, initiate a plan, and manage time and supportive resources needed for assignments.

However, interpreting rubrics was not effortless for students as also seen in our student partnership work cited in Section 2. To better understand the rubrics, students welcomed additional explanations from tutors and worked examples. In the current teaching practice, assessment-support sessions were planned in teaching schedules. There were variations in elaborating assessment criteria, depending on the individual member of staff. For instance, example assignments were provided with detailed annotations or common weaknesses were presented with suggested improvements. The assessment-support sessions resolved student's confusion around how to approach assessments, such as 'using a good range' of academic literature, 'argument', and 'critical' evaluation. Our project findings suggested that the assessment support was particularly important to new students and students who enrolled in novel modules (e.g., neuroscience, work placement) and novel assignments (e.g., reflective portfolio, poster). Insufficient support could lead to assessment anxiety, complaints about awarded grades, and dissatisfactory learning experience.

Although students praised the available resources, they seemed to have difficulty in utilising the rubrics to distinguish between good and poor-quality work. Reflection on peer marking experience suggested that inexperienced students tended to award marks lower than the tutor's mark. This resembles some of the difficulties raised by the Student Panel regarding the perceived usefulness (or lack of) of peer assessment. In terms of their own work, they did not have clear thoughts regarding what was required in a higher-grade band. In addition to providing the breakdown grade descriptors (from excellent to fail) for each assessment criterion, future work could focus on training students to identify distinctive features of each grade band through explicit tutor led activities.

Besides serving the instructional purpose, rubrics are used by markers to evaluate student work (e.g., Campbell, 2005). The grade awarded to each criterion is a way of offering feedback. Such feedback could be dismissed by students due to insufficient details; given that the evaluation was too generic to take any specific practice with a clear aim and scope. Vague feedback was also found to be associated with criticising the fairness of the marking process; differentiating grade boundaries could be perceived subjective and potentially biased. In contrast, students could understand rubrics well within specific contexts. Using the rubric alongside in-text comments supported students in understanding the strengths and weaknesses of their work. For instance, highlighted segments with comments (e.g., text in speech bubbles) associated with specific criterion in the rubric have been praised as very useful feedback. Students could see the evidence for the marker's evaluation and understood what led to the awarded grade. Such understanding increased the likelihood of taking actions (e.g., deliberate practice, seeking additional support/recourses) for improvement.

Although some students actively approached markers to discuss improvements, it was evident that a large proportion of students focused on grades; and had limited interest in engaging with feedback and post-assessment drop-in support. This grade-focused mindset indicated a lack of awareness of long-term

skill growth and the connectedness of modules. To address this issue, our current student partnership project creates student-facing resources with the aim of encouraging a long-term view of learning. It is assumed that explicit identifications of to-be acquired skills and their implications in both academic and career settings could provide a general roadmap of programme learning experience, and students would be inspired to develop their personal provisions of university achievements. Thus, students could become more motivated in engaging with feedback, and thus, achieving better grades, and experiencing higher learning satisfaction.

The creation of the long-term road map of skill development started from mapping skills and knowledge listed in assessment rubrics used in all taught modules. Then our student partner organised the assessed skills in a visual guide demonstrating the pathway of skill growth via modules. The visual guide also directed students' attention to the value of skill development by suggesting application to potential career routes. Moreover, the visual guide highlighted the supportive learning environment (i.e., academic community) that students could seek collaborations and experience belongingness. In the framework of long-term skill development, rubrics present short-term, modular goals that collectively contribute to a transformative learning experience. The completed visual guide will be available to students in companion with the programme handbook in the coming academic year.

Brief Summary

Student voice is crucial in cultivating good learning experience. The student partnership approach to teaching and learning enhancement has been well established in the School. Regarding the use of rubrics, our projects found that students consulted the rubrics both prior to and post assessments for instructions and feedback, respectively. It is essential to provide sufficient explanations and/or examples in line with grade descriptors listed in rubrics. Pieces of information are not perceived as useful feedback unless students can see the context to which the rubric grade descriptors are applied. It is suggested we avoid generic, vague statements in feedback and connect rubric criteria with specific segments of the submitted coursework. In addition to modular expectations, rubrics, like a springboard, could serve as a long-term skill development. We are working on fostering this long-term, holistic view of skill growth to support a transformative learning journey. Further evidence of the importance of long-term road maps is also discussed in Case Study 3.

Case Study 2: Rubrics for Professional Programmes – Improving Assessment Literacy

Defining and measuring competence is crucial and it has been well documented (Hodgson, et al., 2021) that failure to identify and manage impaired competence in applied practitioners can have critical consequences. Therefore, consistent conceptualisation and accurate measurement of competence as a construct is imperative. In psychological therapy training programmes in England, many follow the national training curriculum that aligns with standards of required competence for psychotherapy practice in Improving Access to Psychological Therapy services, which are largely operated by the National Health Service. Although, there is national guidance on rubric construction, there is also flexibility for training providers to adapt rubrics for their own purpose.

At the University of Reading, we deliver several commissioned psychological therapy training programmes, and this case study will focus on our MSci Applied Psychology (Clinical) programme. Our

MSci programme enables students to train as Psychological Wellbeing Practitioners, an accredited role by the British Psychological Society and the British Association for Behavioural and Cognitive Psycho-therapies. Psychological Wellbeing Practitioners work with individuals experiencing common mental health difficulties including depression and anxiety disorders. Therefore, it is imperative that as teachers, we create clear and transparent rubrics that meaningfully support the development of students' clinical skills and ultimately, standards of psychotherapeutic care. In this case study, we start by discussing how we have worked with a student partner to improve student assessment literacy by examining how to create, and disseminate, rubric guidance outside of the classroom e.g., via screencasts and discussion boards on a Virtual Learning Environment. In the second half of this case study, we focus on the importance of improving staff assessment literacy regarding (a) the role of marker training in improving staff awareness, and use, of competency measures and (b) how to deliver good practice rubric-focused teaching and feedback tutorials. The foci of this case study directly complement some of the key considerations raised by the institutional Student Panel in Section 2, namely support for rubric interpretation and improving assessment literacy for all.

Improving Student and Staff Assessment Literacy

We engaged in a student-staff partnership to develop and embed assessment rubric screencasts in the teaching of psychological therapies. The activity was completed due to the complexity of the clinical competency assessments i.e., the clinical competency assessments have many components and so, only providing an in-class overview of the rubrics has some limitations. Therefore, the aim was for students to be able to review screencast tutorials for each of the eight assessments in their own time via Black-board, our Virtual Learning Environment. This was deemed important from an inclusivity perspective as some learners have different processing speeds and some prefer to revisit audio guidance from a lecturer after the information is provided in the initial teaching session. We deemed that brief written guidance on rubrics is not sufficient to significantly improve student assessment literacy or to be inclusive to a range of learners.

The project was completed by Professor Al Laville (Dean for Diversity and Inclusion), Tamara Wiehe (Lecturer in Clinical Psychology) and PhD student partner, Chloe Chessell. It was important to include Chloe in this project as she has lived experience of completing the MSci programme and therefore, could provide unique insight into the benefit of assessment rubric screencasts. The first part of the project involved Chloe assisting in the development of the assessment rubric screencasts as we deem co-production and collaboration to be crucial in developing effective pedagogical approaches. Chloe reflected on the assessment information provided to her as a student to inform what additional material would be useful for current and future student cohorts. This resulted in the screencasts including an audio tutorial covering a) the credit weighting of the assessment and the module it sits in, b) information on the length of the clinical assessment and the overall structure of the assessment, c) guidance on each part of the clinical assessment rubric e.g., interpersonal skills and clinical risk assessment and what is expected to pass, d) where to find assessment exemplars on the Blackboard site to develop student's understanding of what 'good' looks like, and e) how to access further support from lecturers. Chloe's reflections on this activity are as follows:

The screencasts that have been developed added to the information that I had as a student, as this format allows students to review assessment information in their own time, and at their own pace. Screencasts

can also be revisited, which may help students to ensure they have met the marking criteria for a specific assessment. Furthermore, embedded videos/links to information to support the development of key skills within these screencasts expand upon the information my cohort received, and will help students to develop these clinical skills.

The project team agreed that the student-staff partnership was key to the success of the project as we needed to ensure that the student voice was at the forefront. The assessment rubric screencasts have been well received by students as acknowledged in student feedback. We are pleased with the improvements in student's assessment literacy, which is seen in an increased number of students receiving a higher grade in assessments.

To support the assessment rubric screencasts and to further develop student's assessment literacy, we also set-up online discussion boards on Blackboard so students can continue their conversations about assessment rubrics with each other on that platform and ask lecturers for further support as well. To ensure inclusivity was present in this activity, students could send in questions to lecturers over email and receive a direct response. The student's question and the lecturer's answer are then posted on the Blackboard discussion board, with the student's name remaining anonymous. This approach has been highly valued by students and is also reflected in our institutional work on creating inclusive Teaching and Learning environments.

More recently, we completed a student-staff partnership project that explored the importance of improving staff assessment literacy and the relationship this has with providing useful developmental feedback to students (Laville et al., (2023)). In our paper, we recommend that HEIs provide comprehensive training to staff on assessment rubrics as this will improve a) staff assessment literacy, b) staff ability to clearly explain rubrics to students – both in-class and outside of the classroom e.g., assessment rubric screencasts, which will have a positive impact on student assessment literacy, and c) positive impact on the quality of the marking including the quality of feedback provided to students. This approach aims to improve consistency of experience of assessment and feedback for students.

To implement these recommendations, we recommend that institutions provide appropriate marker training at both a module and programme level to ensure that balanced feedback is provided to students for each assignment e.g., a balanced review of 2 strengths and 2 weaknesses in qualitative feedback. This marker training should be delivered by an experienced marker and ideally, a colleague who have been involved in the construction of the assessment rubric. To further support the feedback process, we recommend that institutions provide feedback tutorials to students in order to support students to engage with the feedback on the assessment rubric. These tutorials will function best as one-to-one tutorials based on our research findings and will resolve the engagement with feedback issues that were identified in Case Study 1. Whilst we recognise that these actions require more initial faculty resource, there are realised long-term benefits to the student experience that we have experienced within our own practice. Our findings have been disseminated to several University of Reading groups to support our institutional assessment and feedback strategy.

In summary, by improving staff assessment literacy, faculty members will be better equipped to provide good quality written and verbal feedback to students, which in turn will improve student engagement with assessment rubrics.

Case Study 3: Bridging the Gap — Underpinning Assessments with Generic Rubric*s*

Following on from improving staff assessment literacy and providing feedback tutorials (Laville et al., 2023), this case study provides a more detailed discussion on the use of rubrics as a feedback tool.

The use of rubrics as a method of feedback is common within Higher Education, and although they can be a successful tool, significant differences still exist between staff and student views on feedback (Zhao et al., 2022, Mulliner & Tucker, 2017). Here we present evidence from a student-led study that highlights this disparity and suggests that customisable generic feedback rubrics can increase engagement, improve parity and the student experience, while at the same time addressing concerns of staff. The study highlights the importance of 'closing the loop' dialogue between staff and students to drive change.

The study took place as part of the University's Student Partners Scheme. Four undergraduate students worked in partnership with the School of Biological Sciences to identify and address some of the challenges around feedback from both the staff and student perspective. The project was divided into four stages:

Stage 1: Staff focus group and a student survey designed to identify key challenges.
Stage 2: Students worked to design a generic rubric resource,
Stage 3: The generic rubric resource was evaluated by students and staff.
Stage 4: Students 'closed the loop' by presenting their findings to staff.

The initial survey and focus group supported evidence from the literature of the disparity in the views of students and academics (Mulliner & Tucker, 2017, Zhao et al., 2022). These findings are discussed in detail below.

Staff cited time required to write detailed feedback on scripts (84%), lack of student engagement with feedback (82%), and little evidence that students can feedforward/ transfer skills between modules as their main concerns (78%) and are typical of the views highlighted throughout the literature (Winstone et al., 2017, Zhao et al, 2022). 46% of staff surveyed did not believe that rubrics were an effective way of providing feedback and cited the lack of detail and depth as a major concern.

The biggest concern amongst students was the disparity in feedback quality between modules (cited by 89% of students) while 75% of the students surveyed were concerned about transparency and were not able to see why they had gained, or lost marks compared to another student. The survey gave some insights into some of the concerns of staff in that students did not appear to be able to translate feedback between modules or understand that the same skills were underpinning assessments in different modules. Many of the students surveyed (54%) stated that they did not understand how to improve their weakest areas and admitted that they failed to understand what to do with detailed complex feedback. These views are not new, reflect views expressed earlier in our chapter, and have in the past been linked to lack of engagement with feedback (Sadler, 2010, Adcroft, 2011, Young, 2020). Students surveyed (98%) did however believe that rubrics were the most effective way of providing written feedback.

As a response to this, the student partners in collaboration with two teaching staff created two generic customisable skills-based rubrics (GCR), one for written tasks and one for lab work.

The rubrics had three main foci:

1. To ensure that the same task specific key skills were presented across all modules and years, and these were adapted from those presented to all students in their key skills modules in Part one and two. It was hoped that this would encourage students to translate feedback more effectively between modules and allow them to track their progress across the key skills.
2. To provide 'one click' support for each skill area, this was linked to resources provided by the School and University.
3. To embed flexibility with a customisable element to enable staff to link the rubric to the specific task and add extra marking criteria.

Views (both staff and student) on the GCR were collected using surveys and indicated that the rubrics successfully addressed many of the concerns a raised by both stakeholders. 75% of students believed that generic customisable rubrics would significantly improve their feedback experience (with a further 19% answering maybe). 92% of the students surveyed believed that the rubrics would reduce disparity between modules and 85% felt the one-click skills help would encourage them to engage with their feedback and to link their assignments across modules and improve future work. Interestingly, 90% of students surveyed agreed that combining the rubric with a 'live' feedback session where the lecturer discussed the main points arising, would be an 'ideal' way forward. This mirrors the importance of providing feedback sessions as detailed in Case Study 2.

Students were given the opportunity to make suggestions and comments and it was particularly interesting to see that not a single student highlighted concerns about the loss of detailed feedback written on scripts. In essence the students believed that the rubrics would address both their concerns and those of the staff.

Staff views on the generic rubrics were however less favourable with only 54% of the staff surveyed agreeing that this would be an effective way to provide feedback, and this view was unchanged between the first and second survey. Interestingly, unlike the students many staff continued to cite the loss of detailed feedback on scripts as a concern (something that directly impacts on their other concern about time). There was clearly a need to 'close the loop' and for the staff to 'see' the students' views.

The results of this study were presented to the whole school at a 'Teaching and Learning Away Day'. This allowed the rubric to be presented from the students' point of view and in the context of their concerns. The presentation prompted a wide discussion around feedback, the recognition of the need for more parity and the worth of time-consuming detailed feedback. Following the session, a 'hands up' survey indicated that over 80% of teaching staff within the school were happy to integrate the rubrics into their feedback. This study highlights the importance of a full partnership pathway (including 'closing the loop') in driving change, as when the rubrics were presented in isolation without the discussion, many staff remained resistant to rubrics with their opinions unchanged.

The changes implemented via this initiative could have contributed to the 13.5% increase in the NSS Assessment and Feedback scale across four of our degree programmes between 2021 and 2022.

The results of this study highlight the importance of the partnership between staff and students reported elsewhere in the literature (Handley et al., 2011) to challenge opinions, share ideas, improve engagement and drive effective teaching and learning enhancements. This project is ongoing, and a second set of surveys are planned to explore further future directions in this research.

FUTURE RESEARCH DIRECTIONS

Engaging in student-staff partnership approaches for rubric evaluation has been fundamental and has allowed us to elevate student voice in this area. By engaging in this process, it has become clear that the sector implications for practice are that students require clear guidance on how to differentiate grade boundaries on a rubric and to develop their understanding of what 'good', 'very good' and 'excellent' looks like. This can be supported by providing annotated exemplars, so students can easily identify areas of strength as well as areas for development. Furthermore, students need to be provided with qualitative feedback for each part of the rubric to inform future development strategies. Future research should focus on the effectiveness of the assessment and feedback strategies that we have detailed in Laville et al., (2023), specifically providing balanced feedback and providing meaningful feedback tutorials. This could be realised by collecting qualitative student feedback on the usefulness of balanced assessment feedback and their perceptions of the benefits of engaging in the feedback tutorials as well as realised benefits.

CONCLUSION

Throughout this chapter, we have discussed the benefits of adopting student-staff partnership approaches in the construction, discussion, and evaluation of rubrics. For this approach to be successful, we have been cognisant of the principles underpinning student-staff partnerships as well as the need to further promote the effectiveness of student-staff partnerships across our university community. A significant facilitator in elevating the student voice at UoR has been the Student Panel, which effectively contributes to the development of institutional Teaching and Learning strategy. The latter is ongoing work for us however, we are encouraged to see the increase in student-staff partnerships at UoR and the international dissemination of our projects (e.g., Laville et al., 2023).

From our first case study, we clearly see an implication for practice, which is the need for in-class support to unpack the requirements detailed in the rubric. This in-class support should particularly focus on unpacking the language used in the rubrics and what is expected of students. These discussions should also include how the skills being assessed in the rubric contribute to long-term skill development.

In our second case study, we saw that whilst in-class support for rubrics is very important for improving student's assessment literacy, an implication for practice is that additional support outside of the classroom should also be provided. Laville et al. (2020) shows the importance of providing screencasts that detail rubric criteria and expectations, particularly when focusing on competency measures in professional training programmes. This support is coupled with online discussion boards, so the rubric conversation can be developed outside of the classroom as well. Laville et al. (2023) discusses another implication for practice, which is the importance of also developing the assessment literacy of colleagues via marker training. The marker training delivered by Laville at UoR has developed colleagues understanding of rubrics and thus increased colleagues' confidence in discussing rubrics with students in class. The learning from the marker training has also had a positive impact on marker consistency and thus, an increase in positive student feedback.

In our third case study, we saw the importance of evaluating and revising pre-existing rubrics with students. The creation of the Generic Customisable Rubrics (GCRs) was an innovative approach to addressing the concerns held by students and colleagues about the use of the rubrics, The GCRs directly support the points raised in Case Study 1 by providing rubrics that account for key skills that are present

across all modules and thus, supporting the concept of long-term skill development. The GCRs also support students to plan for future skill development by utilising the 'one-click' function to accessing further resources. This case study illustrated that sector practice needs to focus more on the value and importance of sharing student voice comments with colleagues in order to engage in meaningful discussion about the importance of using GCRs across modules but also providing the option for colleagues to add additional marking criteria too.

In summary, student-staff partnerships are central to the construction, discussion and evaluation of rubrics however, additional support mechanisms such as those detailed above also need to be included to successfully improve student's assessment literacy.

REFERENCES

Adcroft, A. (2011). The mythology of feedback. *Higher Education Research & Development*, *30*(4), 405–419. doi:10.1080/07294360.2010.526096

Advance H. E. (2016) *Student Engagement Through Partnership Framework*. Advance HE.

Advance H. E. (2023), *The UK Engagement Survey (UKES)*. Advance HE. https://www.advance-he.ac.uk/reports-publications-and-resources/student-surveys/uk-engagement-survey-ukes

Andrade, H. (2000). Using Rubrics to Promote Thinking and Learning. *Educational Leadership*, *57*(5), 13–18.

Andrade, H., & Du, Y. (2007). *Student Responses to Criteria Referenced Self-assessment. Educational Administration & Policy Studies Faculty Scholarship*. Scholar Archive. scholarsarchive.library.albany.edu/eaps_fac_scholar/1

Andrade, H. G. (2005). Teaching with rubrics: The good, the bad, and the ugly. *College Teaching*, *53*(1), 27–31. doi:10.3200/CTCH.53.1.27-31

Andrews, M., Brown, R., & Mesher, L. (2018). Engaging students with assessment and feedback: Improving assessment for learning with students as partners. *Practitioner Research in Higher Education Journal*, *11*(1), 32–46.

Arter, J. A., & Chappuis, J. (2007). *Creating & recognizing quality rubrics*. Pearson Merrill Prentice Hall.

Bolton, F. C. (2006). Rubrics and adult learners: Andragogy and assessment. *Assessment Update*, *18*(3), 5–6.

Bovill, C., & Bulley, C. J. (2011). A model of active student participation in curriculum design: exploring desirability and possibility. In C. Rust (Ed.), *Improving Student Learning (18) Global theories and local practices: Institutional, disciplinary and cultural variations* (pp. 176–188). The Oxford Centre for Staff and Learning Development.

Brookhart, S., & Chen, F. (2015). The quality and effectiveness of descriptive rubrics. *Educational Review*, *67*(3), 343–368. doi:10.1080/00131911.2014.929565

Buckley, A. (2021). Crisis? What crisis? Interpreting student feedback on assessment. *Assessment & Evaluation in Higher Education*, *46*(7), 1008–1019. doi:10.1080/02602938.2020.1846015

Campbell, A. (2005). Application of ICT and rubrics to the assessment process where professional judgement is involved: The features of an e-marking tool. *Assessment & Evaluation in Higher Education*, *30*(5), 529–537. doi:10.1080/02602930500187055

Carless, D., & Chan, K. (2017). Managing dialogic use of exemplars. *Assessment & Evaluation in Higher Education*, *42*(6), 930–941. doi:10.1080/02602938.2016.1211246

Cockett, A., & Jackson, C. (2018). The use of assessment rubrics to enhance feedback in higher education: An integrative literature review. *Nurse Education Today, 69*, 8–13.   doi:10.1016/j.nedt.2018.06.022

Cook-Sather, A., Bovill, C., & Felten, P. (2014). *Engaging students as partners in teaching and learning: A guide for faculty*. Jossey-Bass.

Dawson, P. (2017). Assessment rubrics: Towards clearer and more replicable design, research and practice. *Assessment & Evaluation in Higher Education*, *42*(3), 347–360. doi:10.1080/02602938.2015.1111294

Ecclestone, K. (2001). 'I know a 2:1 when I see it': Understanding criteria for degree classifications in franchised university programmes. *Journal of Further and Higher Education*, *25*(3), 301–313. doi:10.1080/03098770126527

Fletcher, R. B., Meyer, L. H., Anderson, H., Johnston, P., & Rees, M. (2012). Faculty and students conceptions of assessment in higher education. *Higher Education*, *64*(1), 119–133. doi:10.100710734-011-9484-1

Giacumo, L. A., & Savenye, W. (2020). Asynchronous discussion forum design to support cognition: Effects of rubrics and instructor prompts on learner's critical thinking, achievement, and satisfaction. *Educational Technology Research and Development*, *68*(1), 37–66. doi:10.100711423-019-09664-5

Graham, A., Harner, C., & Marsham, S. (2022). Can assessment specific marking criteria and electronic comment libraries increase student engagement with assessment and feedback? *Assessment & Evaluation in Higher Education*, *47*(7), 1071–1086. doi:10.1080/02602938.2021.1986468

Handley, K., Price, M., & Millar, J. (2011). Beyond "doing time": Investigating the concept of student engagement with feedback. *Oxford Review of Education*, *37*(4), 543–560. https://www.jstor.org/stable/23047914. doi:10.1080/03054985.2011.604951

Healey, M., Flint, A., & Harrington, K. (2014). *Engagement through partnership: students as partners in learning and teaching in Higher Education*. Academy HE. https://www.heacademy.ac.uk/engagement-through-partnership-students-partners-learning-and-teaching-higher-education

Healey, M., Flint, A., & Harrington, K. (2016). Students As Partners: Reflections on a Conceptual Model. *Teaching & Learning Inquiry*, *4*(2), 8–20. doi:10.20343/teachlearninqu.4.2.3

Healey, M., & Healey, R. L. (2019). *Student engagement through partnership: A guide and update to the Advance HE framework*. Advance HE. mickhealey.co.uk/devtest/wp-content/ uploads/2020/02/SaP-Guide-Updated.pdf

Hodgson, K., Lamport, D., & Laville, A. (2021). Variable trajectory: a systematic review, analytic synthesis and construct domain consolidation of international measures of competence in doctors and medical students. *BMJ Open, 11* (8). doi:10.1136/bmjopen-2020-047395

Jonsson, A. (2014). Rubrics as a way of providing transparency in assessment. *Assessment & Evaluation in Higher Education*, *39*(7), 840–852. doi:10.1080/02602938.2013.875117

Laville, A., Chessell, C., & Wiehe, T. (n.d.) Developing and embedding electronic assessment overviews. Teaching and Learning Exchange, CQSD: University of Reading.

Laville, A., Holtom, H., Conway, E., & Alder, C. (2023). Exploring clinical perfectionism in higher education students: key recommendations and reflections on a partnership. *International Journal for Students as Partners, 7* (1).

Lorber, P., Rooney, S., & Van Der Enden, M. (2019). Making assessment accessible: A student–staff partnership perspective. *Higher Education Pedagogies*, *4*(1), 488–502. doi:10.1080/23752696.2019.1 695524

Martens, S. E., Spruijt, A., Wolfhagen, I. H. A. P., Whittingham, J. R. D., & Dolmans, D. H. J. M. (2019). A students' take on student-staff partnerships: Experiences and preferences. *Assessment & Evaluation in Higher Education*, *44*(6), 910–919. doi:10.1080/02602938.2018.1546374

Matshedisho, K. (2020). Straddling rows and columns: Students' (mis)conceptions of an assessment rubric. *Assessment & Evaluation in Higher Education*, *45*(2), 169–179. doi:10.1080/02602938.2019.1616671

Mercer-Mapstone, L., & Bovill, C. (2020). Equity and diversity in institutional approaches to student–staff partnership schemes in higher education. *Studies in Higher Education*, *45*(12), 2541–2557. doi:1 0.1080/03075079.2019.1620721

Mercer-Mapstone, L., & Marie, J. (2019). *Practical guide: Scaling up student-staff partnerships in higher education*. Institute for Academic Development: University of Edinburgh. http://bit.ly/2EfUR16

Moys, J.-L., Collier, J., & Joyce, D. (2018). By design: engaging Graphic Communication students in curriculum development. *The Journal of Educational Innovation, Partnership and Change, 4* (1). https://centaur.reading.ac.uk/76158/ doi:10.21100/jeipc.v4i1.752

Mulliner, E., & Tucker, M. (2017). Feedback on feedback practice: Perceptions of students and academics. *Assessment & Evaluation in Higher Education*, *42*(2), 266–288. doi:10.1080/02602938.2015.1103365

Norton, L. (2004). Using assessment criteria as learning criteria: A case study in psychology. *Assessment & Evaluation in Higher Education*, *29*(6), 687–702. doi:10.1080/0260293042000227236

Ntem, A., Nguyen, E., Rafferty, C., Kwan, C., & Benlahcene, A. (2020). Students as partners in crisis? Student co-editors' perspectives on COVID-19, values, and the shift to virtual spaces. *International Journal for Students as Partners*, *4*(2), 1–8. doi:10.15173/ijsap.v4i2.4432

NUS. (2015). *Assessment and feedback benchmarking tool.* QAA. https://www.qaa.ac.uk/docs/qaas/focus-on/nus-assessment-and-feedback-benchmarking-tool.pdf?sfvrsn=f37cf481_14

O'donovan, B., Price, M., & Rust, C. (2001). The Student Experience of Criterion-Referenced Assessment (Through the Introduction of a Common Criteria Assessment Grid). *Innovations in Education and Teaching International*, *38*(1), 74–85. doi:10.1080/147032901300002873

O'donovan, B., Price, M., & Rust, C. (2004). Know what I mean? Enhancing student understanding of assessment standards and criteria. *Teaching in Higher Education*, *9*(3), 325–335. doi:10.1080/1356251042000216642

O'donovan, B., Price, M., & Rust, C. (2008). Developing student understanding of assessment standards: A nested hierarchy of approaches. *Teaching in Higher Education*, *13*(2), 205–217. doi:10.1080/13562510801923344

OfS. (2020). *Teaching Excellence Framework*. OfS. https://www.officeforstudents.org.uk/advice-and-guidance/teaching/about-the-tef/

OfS. (2022). *Building a culture of student engagement: our priorities for 2022-23, Student Engagement Strategy*. Available from https://www.officeforstudents.org.uk/about/student-engagement-strategy/

OfS. (2023). *Methodology document for National Student Survey sector analysis.* OfS. https://www.officeforstudents.org.uk/media/de0d31cf-23d7-42c7-8113-500aefcc96fb/nss-2022-sector-analysis-method-document.pdf

OfS. (2023). *National Student Survey (NSS)*. OfS. https://www.thestudentsurvey.com/

Padden, L., & O'Neill, G. (2021). Embedding equity and inclusion in higher education assessment strategies: Creating and sustaining positive change in the post-pandemic era. In P. Baughan (Ed.), *Assessment and feedback in a post-pandemic era: A time for learning and inclusion* (pp. 138–147). Advance HE. https://www.advance-he.ac.uk/knowledge-hub/assessment-and-feedback-post-pandemic-era-time-learning-and-inclusion

QAA. (2014). *UK Quality Code for Higher Education Part A: Setting and Maintaining Academic Standards PART A The Frameworks for Higher Education Qualifications of UK Degree-Awarding Bodies.* QAA. https://www.qaa.ac.uk/docs/qaa/quality-code/qualifications-frameworks.pdf

QAA. (2018). *UK Quality Code, Advice and Guidance: Student Engagement.* QAA. https://www.qaa.ac.uk/the-quality-code/advice-and-guidance/student-engagement#

Reddy, Y., & Andrade, H. (2010). A review of rubric use in higher education. *Assessment & Evaluation in Higher Education*, *35*(4), 435–448. doi:10.1080/02602930902862859

Sadler, D. (2009). Indeterminacy in the use of pre-set criteria for assessment and grading. *Assessment & Evaluation in Higher Education*, *34*(2), 159–179. doi:10.1080/02602930801956059

Sadler, D. R. (2010). Beyond feedback: Developing student capability in complex appraisal. *Assessment & Evaluation in Higher Education*, *35*(5), 535–550. doi:10.1080/02602930903541015

Sambell, K., & McDowell, L. (1998). The construction of the hidden curriculum: Messages and meanings in the assessment of student learning. *Assessment & Evaluation in Higher Education*, *23*(4), 391–402. doi:10.1080/0260293980230406

Student Minds. (2019). *The University Mental Health Charter Framework*. University Mental Health Charter. https://universitymentalhealthcharter.org.uk/themes/

University of Reading. (2018). *Marking and Feedback.* University of Reading. https://www.reading.ac.uk/cqsd/-/media/project/functions/cqsd/documents/qap/10-marking-withannexes.pdf?la=en&hash=AF70ED315F16D1ACC0322306C53138DE

University of Reading. (2020). *Partnership at UoR Guide.* University of Reading. https://sites.reading.ac.uk/wp-content/uploads/sites/35/2020/10/Partnership-at-UoR-Guide-October-2020.pdf

University of Reading. (2021). *Curriculum Framework*. University of Reading. https://www.reading.ac.uk/cqsd/-/media/project/functions/cqsd/documents/qap/university-of-reading-curriculum-framework.pdf

Winstone, N. E., Nash, R. A., Parker, M., & Rowntree, J. (2017). Supporting Learners' Agentic Engagement With Feedback: A Systematic Review and a Taxonomy of Recipience Processes. *Educational Psychologist*, *52*(1), 17–37. doi:10.1080/00461520.2016.1207538

Zhao, X., Cox, A., Lu, A., & Alsuhaibani, A. (2022). A comparison of student and staff perceptions and feelings about assessment and feedback using cartoon annotation. *Journal of Further and Higher Education*, *46*(5), 586–604. doi:10.1080/0309877X.2021.1986620

KEY TERMS AND DEFINITIONS

Academic Representation: The Course Representative scheme, run by the Students' Union, embeds student voice into quality assurance process at the University. Course Reps are student volunteers who represent and report the views of their student cohort on teaching and learning matters.

Assessment Literacy: The process of developing insight into the meaning of specific assessment criteria.

Generic Customisable Skill-Based Rubrics: Rubrics that support long-term learning and development by providing consistency of skill identification across assessment types

Partnership: The process by which individuals work collaboratively, bringing together the perspectives, knowledge and skills of students and staff, to achieve a shared goal.

Student Panel: The name given to the scheme that employs 50 students to actively contribute to shaping the direction of strategic projects at the University of Reading.

Student Partner: The role given to a student formally or informally engaged in collaborative work with a member(s) of staff, and/or students at the University of Reading

Student Voice: The collective term used to describe the views and perspectives of individual students and the wider student population in the teaching and learning context.

Chapter 8
Co–Production of Assessment Rubrics in an Online Education Context

Anja Harrison
King's College London, UK

Maren Breier
King's College London, UK

Harriet Power
King's College London, UK

Brenda P. Williams
King's College London, UK

ABSTRACT

Marking rubrics is hailed as a transparent and effective way of supporting student success; enhancing their ability to understand and use their learning environment to achieve their goals. Rubrics also enable staff to mark fairly and consistently. Yet, to be successful, rubrics must be understandable to all and there needs to be active engagement from students and staff alike. Understandability requires that the wording be inclusive and considerate of student diversity, and this is especially true when considering online courses where the student body is often more culturally diverse. Co-creation with students can promote inclusivity and the development of meaningful and successful rubrics. This chapter provides a step-by-step guide for co-creating and implementing rubrics in an online education context, developed through collaboration with the co-creation student panel from the online programmes at the Institute of Psychiatry, Psychology and Neuroscience, KCL.

THE DEFINITION OF A RUBRIC

The word rubric can mean different things in different contexts. In a broad sense, a rubric is a collection of statements that simultaneously describe criteria and measure quality. In education, a rubric refers to a scoring tool composed of a set of criteria/descriptors and an accompanying scale illustrative of differing levels of student performance. A rubric is not an assessment brief, or the assessment criteria themselves. An assessment brief is the set of instructions provided for students that outlines what is required and expected of them to complete the task. In some instances, this may include information about the assessment criteria, which will inform the rubric and final grades, however, this is not always the case. The assessment criteria outline what knowledge and skills a student should be able to demonstrate on completion of the assessment. Assessment criteria are aligned with course and module learning outcomes and are to be used in the context of assessment standards, which usually include grade descriptors at a more general institutional level (Sadler, 2005). A rubric then links the assessment criteria to performance indicators, providing a series of statements that describe the quality of learning (Brookhart, 2018) in a clear scheme for scoring (Reddy & Andrade, 2010). Effective rubrics can support both staff and students by facilitating more expeditious marking and providing more consistency (Jonsson and Svingby, 2007; Brookhart and Chen, 2015; Panadero & Jonsson, 2013). This is achieved by providing clear expectations for students on how their performance will be assessed, together with the opportunity to measure the quality of their own work; whilst for staff they provide a framework that has been shown to reduce the time needed to evaluate a piece of work (Stevens & Levi, 2013) and contribute to consistency of grading across both students and modules. Yet, for rubrics to have a true impact they must be understood and successfully applied by all stakeholders, and co-creation can be key in achieving this.

The Importance of Co-Creating Rubrics

Assessments enable educators to evaluate the potential and actual achievements of students, and the linked feedback serves as a basis for enabling students to improve their future performance (Williams, 2000). Good assessment performance and meeting learning outcomes requires students to engage with materials and understand what is expected of them. However, there is a delicate balance between providing too little information around what needs doing to perform well and there being too detailed or prescriptive information made available that will prevent in-depth learning and give assessment more of a tick-box feel (Torrance, 2007). Rubrics should not limit or constrain students in what their assessments can include but give them clarity with appropriate degrees of freedom. To ensure these rubrics are pitched at the right level and optimise education experience, students need to be and feel heard (Murdoch et al., 2020) through engagement and interaction (Peacock et al., 2020; Kahu & Nelson, 2018). The aim is to prevent students from taking a passive role in their education, where they only do what the educator might expect of them to gain a good grade (Reynolds-Keefer, 2010). Co-creation moves beyond teaching students how to use rubrics, which again can result in them simply seeing it as a tool to gain the best mark, rather it is about inviting students' input in the creation of accessible and useful rubrics, making them an integral part of the design process. Evidence suggests such co-creation can have a positive impact on motivation, self-efficacy and hence depth of learning (Arter & McTighe, 2001, Andrade & Du, 2005; Andrade et al., 2009, Reynolds & Keefer, 2010) and, in this way, become an integral part of learning through reflection (Fraile et al., 2017) leading to students' and educators' common goal: better student performance. Moreover, increased belief in one's own ability to learn and complete tasks has

been shown to have a positive effect on a person's education (Linnenbrinch & Pintrich, 2003). However, rubrics are only truly effective at promoting self-assessment if the curriculum allows students to use this understanding to improve their work. This means using rubrics in formative assessments and/or in summative assessments that feedforward (Zimmerman, 2000, Reddy & Andrade, 2010).

Preconditions to Co-Creating Good Rubrics

To create rubrics that are fit for purpose, several factors need to be considered. First, the type of assessment (e.g., oral, written, infographic) needs to be considered when drafting the assessed criteria (Ulker, 2017). Assessment should be authentic and enable students to develop skills relevant for their future career (Wiggins, 1990). As such, rubrics should reflect this and measure these transferable skills. Second, the rubric needs to clearly map onto the broader learning outcomes of the module or programme, that is, criteria need to be appropriate (Brookhart, 2018). Third, a useful rubric should clearly describe the different levels of mastery and all characteristics that are going to be rated as part of the marking process (Allen & Tanner, 2006). Fourth, rubrics need to be designed with inclusivity in mind and a clear understanding of cultural differences that may impact on the interpretation of them (Mills, 2022). Finally, as mentioned above, rubrics need to allow for individualised feedback / feedforward to maximise usefulness and enhance learning (Montgomery, 2002, Zimmerman, 2000, Reddy & Andrade, 2010).

Co-Creation and Learning in an Online Environment

Inclusivity, student engagement and self-efficacy are central goals in all learning environments and co-creation can help facilitate these (Fraile et al., 2017). However, their importance is heightened in the online space where diversity, distance from staff and fellow students are central themes, and self-directed learning is a necessity. Moreover, the inclusive nature of co-creation has the potential to overcome student misinterpretation of requirements for assessments, often evident when a student body is culturally and educationally diverse, as in an online environment (Tanis, 2020, Martin & Bolliger, 2018, Voltmer et al., 2022). Considering diversity is essential when developing rubrics, to ensure that the vocabulary used is clear to all and potential conflicts with a student's cultural experiences are considered (Lewis, 2021). Co-creation with students can help avoid these issues. In addition to student diversity, online education is often characterised by asynchronous interactions and perceived distance (Bolliger & Halupa, 2018). These barriers can be overcome by embedding active student-teacher interaction and collaboration in the course structure (Bolliger & Halupa, 2018). Co-creation of rubrics can facilitate students' inclusion and engagement, which have been shown to contribute to students' success in an online environment (Bolliger & Halupa, 2018).

Most often, the set-up of online learning also requires students to work autonomously and auto-didactically (Chaker et al., 2022). This places a high responsibility on both the student and the teacher to ensure students are well equipped to manage, pace, and advance their own learning and performance. Rubrics can be a critical support for the student and the instructor to facilitate and guide self-paced training. Effective rubrics enable students to monitor and direct their own learning (Panadero et al., 2017). Student ownership of their academic performance has been associated with improved learning outcomes and can be enhanced by involving students in the creation of rubrics (Fraile et al., 2017; Panadero & Jonsson, 2013).

The goal to successfully co-create and implement rubrics in an online environment also responds to challenges posed by the current rise in online higher education (particularly accelerated by the Covid-19 pandemic) (Carroll & Conboy, 2020; Wood, 2017) and the transformation within academia towards a student-centred approach, which acknowledges that engaging students in the assessment process enhances self-awareness and thereby improves students' learning outcomes (Fraile et al., 2017).

With much evidence endorsing rubric co-creation as an important intervention for promoting deep and reflective learning as well as inclusivity, the next question is how students and staff can engage in this process to deliver a favourable outcome for all. One way forward is through using Normalisation Process Theory.

Using Normalisation Process Theory (NPT) as a Guiding Model

Normalisation Process Theory (NPT) is a conceptual framework aimed towards facilitating both the development as well as the integration of complex interventions into routine and already established processes (Murray et al., 2010). It focusses on the practical reality of an intervention determined by individual and collective action and behaviour (May et al., 2018). NPT uses four sociological mechanisms to measure how likely an intervention is to become normalised or integrated into regular practice. These mechanisms look at how well the intervention makes sense to those involved (coherence), how engaged individuals are in the intervention (cognitive participation), how individuals work together to implement the intervention (collective action), and how individuals reflect on and adjust the intervention as needed (reflexive monitoring). These mechanisms help us understand how people interact with and adopt interventions in real-world settings (May et al., 2009).

Through collaboration with a co-production student panel, NPT has been used as the conceptual model for how to co-create and implement rubrics in an online education context. This co-production (co-pro) panel of postgraduate taught online students was set up to engage with the online programme team at the Institute of Psychiatry, Psychology and Neuroscience (IoPPN), in focused discussions around new initiatives. The co-creation of rubrics was one such initiative. Such collaboration with students aligns with King's College London's 2020-25 five-year strategy that states, 'we want students to be co-constructors of the curriculum, and we will incorporate their opinions and suggestions in curriculum design and development'. The panel format of student-staff engagement is embedded at College level where we have a cross-faculty student panel called the King's 100 (https://www.kcl.ac.uk/students/finding-my-voice-building-connections-and-shaping-the-student-experience-at-kings) and at Faculty level, for example, the IoPPN 25, which is a cross programme PGT student panel reflective of the IoPPN student community. These initiatives enable students to work with staff to shape their experience at KCL. As such, the setting up of a co-pro panel for our online students was considered an effective way to hear the student voice where a large student cohort is involved. However, there was a concern about self-selection bias as the aim was for this student group to represent a diverse online student body. To help avoid this issue, a 'recruitment' call went out to all students studying on the online MSc programmes 'Psychology and Neuroscience of Mental Health' and 'Applied Neuroscience'. To apply, students were requested to complete an online form, including their programme of study, a short paragraph on their motivation for applying and on what experience they had of representing others (Appendix 1). The form was developed through conversation with our Faculty Student Experience Manager and Faculty Diversity and Inclusion Lead. The process of selection involved several key considerations to ensure representation and diversity. Here are the steps taken to assemble the panel:

1. Open-ended Questions: To create an inclusive and welcoming environment, the initial step was to keep the questions as open as possible. This allowed students to express any protected characteristics they have, should they wish to disclose them. By providing this opportunity, the aim was to ensure that no individual felt excluded or marginalized based on their background.

2. Advert Promoting Diversity: The advertisement for panel recruitment explicitly highlighted the commitment to creating a representative panel. It emphasized the importance of diverse perspectives and encouraged individuals from different backgrounds to apply. By openly stating the intention to foster inclusivity, the aim was to attract applicants from a wide range of communities and experiences.

3. Equal Opportunity-Type Monitoring Questions: During the selection process, questions were asked in a format similar to equal opportunity monitoring questions. These questions were designed to collect information about the applicants' backgrounds, not as a means of discriminatory selection but rather to assess the overall diversity of the panel. It was made clear to the applicants that this information was gathered with the purpose of ensuring representation of the community in the panel.

4. Ensuring Representation without Quotas: It is essential to be mindful that selection cannot be based solely on protected characteristics, as this would be a violation of the Equality Act. However, collecting information on these characteristics helps understand the current level of diversity within the panel and guides efforts to represent various backgrounds more effectively in the future. It is important to note that having a quota, which would require a specific number or percentage of individuals from certain backgrounds, is not legally permissible.

By following these steps, the aim was to create a co-pro panel that genuinely represented our online community and included a diverse range of perspectives.

As selection purely based on protected characteristics was not possible (see points 3 and 4 above) the final panel of 31 was composed of students from all stages of their online study journey, representing diverse ethnicities and nationalities, a ratio of 2:1 female to male (reflective of programme composition), and those that had experience of representing others and those that didn't in equal numbers. Feedback suggested that students saw the co-pro panel as an excellent opportunity to engage and interact with other students and staff alike. Offering the chance to build connections that would otherwise be difficult to establish given the remote nature of the online programmes and providing a promising initiative that allowed them to actively participate in shaping the student experience.

Discussion and debate among staff and the co-pro panel resulted in the process summarised in Figure 1, incorporating NPT's four components (coherence, cognitive participation, collective action, and reflexive monitoring). This approach ensured practicality and dynamicity as data from NPT analysis constantly informs the intervention process at all stages; design, implementation, embedding and integration (May et al., 2018). In the sections below, each of the four components of NPT will be discussed in turn in relation to the co-created guide for rubric implementation (Figure 1).

Figure 1. Strategy for co-producing rubrics using normalisation process theory as a guide
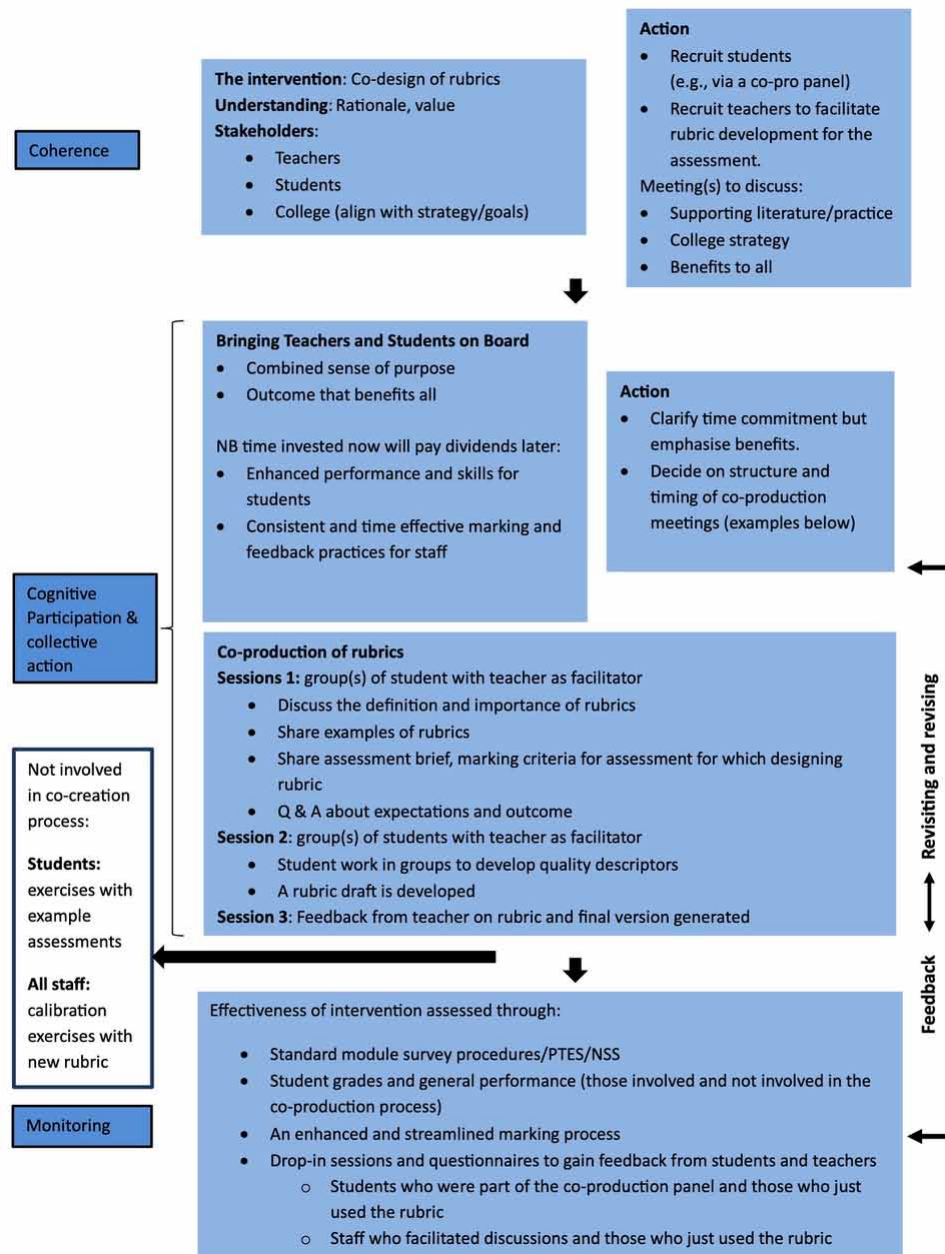
Coherence

For any new practice to be successful and sustainable, there needs to be a clear and coherent understanding among all participants about the nature of the innovation, the work it entails, and the intended outcome and benefits (Wood, 2017). Such sense-making is strengthened/supported by defining a shared purpose at the beginning of the intervention which will then guide the following process (May et al., 2022).

Rubrics are supposed to be 'recipes for success', helping increase transparency of marking criteria for students and teaching staff alike (Jonsson, 2014; McNair, 2016; Allen & Tanner, 2006). Reliability and validity of awarded grades is crucial and can be achieved more easily when using well-designed rubrics (Jonsson & Svingby, 2007), so long as markers are well trained in how to use them (Rezaei & Lovorn, 2010). However, successful use and adaption of rubrics by both students and staff are subject to 'buy-in'. Co-creation has been known to enhance 'buy-in' (Hughes et al., 2019; Snelling et al., 2019), which is also an important step towards 'normalisation', or routine implementation, of new initiatives in many contexts, including educational organisations (Wood, 2017). Rubrics are supposed to enhance self-regulated learning and help students understand and work more effectively with feedback (Jönsson & Panadero, 2017), which requires active engagement and a willingness to carry out self-assessment (Andrade & Boulay, 2003). Co-creation, therefore, presents itself as a suitable tool for creating meaningful rubrics and simultaneously enhancing engagement and chances of permanent implementation of rubric use (Fraile et al., 2017; Panadero et al., 2017).

While designing the intervention, the gap between theoretical concept and reality needs to be considered (Finch et al., 2012). Many external influences impact individual and collective behaviour and thus define the success of the intervention (Finch et al., 2012). Such factors include time constraints, the need to find consensus and compromises between all participants involved, and institutional limitations given that co-creation of rubrics is subject to academic guidelines. In addition, the division of power is key for rubric co-creation to be successful. Any hierarchies among participants need to be defined clearly at the start, indicating the role, responsibility, and authority of everyone involved. This is only possible if participants agree on a shared purpose at the very beginning. Here, the guiding purpose is to enhance the students' learning experience and outcome as well as the teachers' ability to facilitate learning.

Cognitive Participation and Collective Action

Conventionally, rubrics were created by educators and predominantly used by them (Kilgour et al., 2020) and their purpose was mainly to maximise speed and consistency; however, as discussed, they present the opportunity for more and serve to shift attention towards the learner. Indeed, inclusion at the stage of rubric creation results in a greater understanding on the part of both the student and the teacher (Kilgour et al., 2022). In co-creation, students and teachers co-produce a rubric they then go on to use collaboratively. Co-ownership changes the students' perception of rubrics, allowing for a deeper approach to learning beyond maximisation of grade. However, for co-creation to work both teachers and students must have a shared vision; that this process is an important way to promote confidence and self-regulation in students, while enhancing and streamlining marking and feedback practices (Arter & McTighe, 2001; Fraile et al., 2017; Stevens & Levi, 2013). Even with a shared vision there are challenges to the co-creation process and some of these together with potential solutions are listed in Table 1.

For true co-creation students should be actively involved in developing the rubric from scratch. Providing an already developed rubric for an assessment and asking for comment is not co-creation. Students, however, may have very different ideas about rubrics and their purpose so an initial discussion outlining the aims of the collaboration is essential. For online programme staff and the co-production panel at the IoPPN, this included a discussion of the literature explaining the benefits of rubrics (for example, Ragupathi & Lee, 2020) together with creating that shared vision; enhancing the students' learning experience and outcome as well as the teachers' ability to facilitate learning.

Table 1. Challenges and possible solutions when creating rubrics

Challenge	Solution
Co-creation is time consuming.	• Set expectations and goals from outset. • Agree a meeting schedule (virtual for online students), number of meetings required. • Create that shared understanding of a positive outcome in terms of quality of education for staff and students.
Students not understanding assessment language.	• Minimise use of meta language when discussing and developing rubrics, making them more accessible for all. • Create a glossary to support understanding of assessment terminology (see Table 3).
Disagreement among Teachers & Students.	• Promote open dialogue. • Discuss purpose of the rubric and best practices for using them (from perspective of student and teacher). • Link discussion back to learning outcomes, Institution marking criteria and Institutional strategy/goals.
Sharing resources, ideas and promoting interactions.	• Ensure adequate resources available and that they are shared in a timely fashion (e.g., rubric examples, assessment criteria and guidance documents). • Have a central place where all documents can be shared. • Provide adequate opportunities for interactions between meetings (e.g., chat functions via MS Teams). Opportunity for asynchronous contribution keeps participation levels high.

The intended benefits of the intervention were clearly defined as improving communication and collaboration between students and teachers, fostering the development of transferable and metacognitive skills in students, as well as allowing for a more transparent, effective, and student-centred marking process. Such discussions facilitated coherence or 'sense-making' among students and staff providing a clear understanding of the intervention as well as its purpose (May et al., 2022).

An effective next step involved sharing and discussing the learning outcomes of the programme and/or module, assessment guidance and marking criteria and discussing how to use these to create a rubric, introducing exemplars of current rubrics to support this discussion (see Skillings & Ferrell, 2000). Discussion of exemplars also had the advantage of highlighting how currently used rubrics could be improved. The student panel received various rubric templates proposed by staff, as well as the respective learning outcomes and assessment criteria. Each member of the panel gave their feedback in a shared document before discussing it in an online meeting together with staff and panel students. A combined understanding of rubrics was developing and some important insights for improvement of current rubrics was obtained (Table 2). Following these initial meetings, the students then generate the performance statements that link to each grade band with the teacher providing feedback and facilitating discussion.

This 'development' stage can be approached in several ways: rubrics developed through the setting up of a single staff-student group; rubrics developed separately by different student groups each facilitated by a teacher, who then compare and contrast their statements before creating together a final draft for feedback; in groups where each group works on statements for a different grade, which are then combined, discussed and refined; where students work separately to develop their performance statements and then come together for discussing drafting.

Table 2. Feedback from co-production panel on current rubrics

Assessment type for which rubric designed	Student feedback
Mock grant proposal	Student commented on the chosen rubric categories: 'I really like the listed criteria of this rubric. Maybe an additional criterion relating to literature/references could be included: How well does the proposal integrate current and relevant literature and how well does it identify and address gaps in current literature.'
Infographic	This rubric included the category 'Content -Accuracy' in which the highest score stipulated 'at least 4 accurate facts/concepts are displayed in the infographic.' Student suggested: 'I am not sure if the emphasis should be placed on the number of facts, or rather on the quality of the facts presented, how they are presented, how well they are linked and contribute to the flow of information presented, and how easily the reader understands the key facts as main take-aways.'

The most appropriate way forward should be decided through discussion at the first meeting. At the IoPPN a single group approach was used. Throughout this process it is important that the teacher be open-minded, guiding students through a process of debate and negotiation, yet allowing everyone to share their thinking (Kilgour et al., 2020). This is especially true when, for pedagogic reasons, the teacher considers that a performance statement presented by students should not be included or needs substantial revision. Overall, it is important that students feel heard and never undermined. In this way, students become invested in the process, learn about what is important to succeed and how to work as an effective team. At the end of the process something new will have been created and written in a style that everyone can understand and use to enhance self-directed learning (Reddy & Andrade, 2010), promote assessment for learning principles (Panadero & Jonsson, 2013) and equity (Ragupathi & Lee, 2020).

Once rubrics have been created, both students and staff can use these in exercises where strengths and weaknesses of example assessments are discussed, a process that has also been shown to enhance understanding and performance (Andrade et al., 2009). This could be an effective way of developing an understanding of rubrics across the whole student cohort. Regardless, if these co-created rubrics are developed using language that is clear to all students, they should also enable those not involved in the co-creation process or exemplar exercises to self-regulate and enhance their performance (Fraile et al., 2017). To further aid students in understanding the language used in assessment, particularly if not involved in the co-creation process, a glossary could be developed to work in parallel with the rubrics used. An example of such a glossary developed in collaboration with the IoPPN co-pro panel is shown in Table 3. Ultimately, as students see the worth of rubrics to support their learning, they will champion their use by passing on their understanding of rubrics to fellow students. If staff who are not involved in the rubric creation are required to use the rubric, marking calibration exercises should take place ahead of the marking process (see reflexive monitoring). Following the scheduling of sessions with example assessments for students and marking calibration exercises for staff, the rubric is ready to be used and its effectiveness assessed.

Reflexive Monitoring

The effectiveness of the co-creation approach can be monitored and evidenced by the improvement in student outcomes, engagement, and confidence of students, as assessed through established module feedback processes, maybe even the Postgraduate Taught Experience Survey (PTES) or National Stu-

dent Survey (NSS). It is important to determine whether an enhanced understanding and performance is evident for student both involved and not involved in the co-creation process (Fraile et al., 2017). If this is not the case, then consideration could be given to involving all students in the co-creation process (if class size allows) or expanding sessions where the rubric is used to evaluate the quality of example assessments (Andrade et al., 2009). As new students join the programme, example sessions can be held to promote understanding on the importance of rubrics for learning. These sessions may reveal opportunities to further enhance clarity and improve the rubric (Table 2).

Assessing the effectiveness of the rubric for staff, beyond those involved in the co-creation process, can be as simple as asking them whether what they feel would be a fair mark for the piece of work aligns with the mark outcome when the rubric is applied, as well as whether the process is now more streamlined and sustainable. The caveat here is that it may take time for staff to get used to applying a new rubric if they were not involved in the co-creation process. This problem can be alleviated through a programmes standard marking calibration exercises (Sadler, 2011; Sadler, 2013; Watty et al., 2013) where applying the rubric can be discussed and inter-rater reliability tested. If the rubric is effective, when marking the same set of assessments all markers should arrive at an equivalent score.

As a curriculum is revised, the effectiveness of assessments and their associated rubrics should be revaluated to make sure that the rubrics align with current learning outcomes, effectively evaluate and discriminate the quality of student assignments, and can be consistently applied by markers. This could involve organising further discussions with the student body or a co-pro panel. When new assessments are developed then the co-creation process can be revisited to make sure that again the student voice is heard and that rubrics remain a 'recipe for success' for both staff and students.

Table 3. Examples of commonly used 'educational jargon' that may be interpreted differently by people, limiting the quality and usefulness of any rubrics

Commonly used Jargon	Explanation
Avoid secondary references	Cite the original reference rather than refer to a secondary article ('as cited in X et al., 2000') to avoid misrepresentation.
Avoid direct quotes	Demonstrate understanding by explaining and presenting arguments using your own words. It is impossible to determine level/ depth of understanding based on direct quotes.
Assess the limitations of the study	Weigh up aspects of the study and consider weaknesses that might undermine its validity and/or suggest ways the research could be improved. The weaknesses could be methodological or how the authors interpret and present their own findings.
Balanced argument	While it is often valuable to take a stance, be sure to present evidence for the other sides of the argument. Present both arguments and counterarguments to then weigh up and inform your own argument.
Clarity	Make sure the reader can easily understand what points you are making by writing clearly and explaining why you have made these points. This may just be a case of writing straightforwardly, and not assuming the reader will know what you were thinking.
Critically Evaluate/ Critically Analyse	Actively think about and question the claims you are describing or making. Even if the claims are completely valid, show that you haven't just accepted them at face value.
Discuss rather than describe/list	Do not provide just a descriptive account of a study's findings. Compare and contrast the evidence.
Depth/Elaborate/Demonstrate your understanding	Explain arguments in detail, using examples and working through ideas rather than simply glossing over them.
Explain link to question	Make it clear how the information is relevant for answering the question asked.
Flair	Showing a sophisticated or elegant writing style or presenting evidence in an original and insightful way.
Flow	Creating a coherent argument by connecting points in a logical order to ensure that the work is easy to follow.
Focus	Include only material that is directly relevant for answering the question asked.
Illustrate	Give examples to back up the points made, using evidence.
Originality	Demonstrating your own thinking, perhaps by drawing upon research beyond the ones you learned about in class, to make an argument that not every student would have thought of.
Range of material	Avoid basing too much of your work on a couple of references.
Relevance	The information presented does not answer the question asked.
Structure	A way of presenting your work so the reader can follow the argument. Make sure your paragraphs are in a logical order, that you show the connections between different paragraphs, and that each section has good beginning and ending sentences.
Synthesis/Integration	Show how different sources and theories go together to make a good argument.
Unsubstantiated Claims	An unsubstantiated claim lacks evidence. Justify any argument by supporting each point with empirical evidence and references. This will create a more persuasive argument.
Vague	Being too vague about a point by not explaining it in specific language, or by failing to ground it in theory or to use examples.

CONCLUSION

Co-creating rubrics together with students addresses many changes and challenges within current assessment processes in academia/higher education (Fraile et al., 2017; Wood, 2017). While rubric co-creation

has been shown to enhance the quality of marking and feedback procedures for both students and teachers, the benefits reach beyond the grading itself (Kilgour et al., 2020; Ragupathi & Lee, 2020). Involving students in the development of rubrics fosters students' ownership of their academic progress, which has been associated with an improvement in their learning experience as well as outcome (Fraile et al., 2017; Panadero & Jonsson, 2013). However, implementing effective rubric co-creation within universities is subject to various constraints as it requires both structural as well as behavioural changes within the academic set-up (Finch et al., 2012). NPT responds to this reality by serving as a practical guide and toolkit to support academic institutions in embedding new initiatives, such as rubric co-creation, into their existing (marking) procedures (May et al., 2009; Murray et al., 2010). At the IoPPN, co-creation of rubrics within the NPT framework was successfully initiated by involving the IoPPN's student co-pro panel to define a shared purpose which is now guiding the co-development of effective ways to integrate the student voice in assessment processes. As one of its key advantages, NPT allows for consistently incorporating students' and teachers' feedback into the process to further improve the co-creation itself as well as its application into wider academic routine (May et al., 2022). However, establishing rubric co-creation as a new standard will be dependent on buy-in from students and staff alike. This can only be accomplished by raising awareness of its benefits as well as collecting data on its application to generate solid empirical evidence within academia (Kilgour et al., 2022).

ACKNOWLEDGMENT

REFERENCES

Adarkwah, M. A. (2021). The power of assessment feedback in teaching and learning: A narrative review and synthesis of the literature. *SN Social Sciences*, *1*(3), 75. doi:10.100743545-021-00086-w

Allen, D., & Tanner, K. (2006). Rubrics: Tools for Making Learning Goals and Evaluation Criteria Explicit for Both Teachers and Learners. *CBE Life Sciences Education*, *5*(3), 197–203. doi:10.1187/cbe.06-06-0168 PMID:17012210

Andrade, H., & Du, Y. (2005). Student Perspectives on Rubric-Referenced Assessment. *Educational & Counseling Psychology Faculty Scholarship*. https://scholarsarchive.library.albany.edu/edpsych_fac_scholar/2

Andrade, H., & Valtcheva, A. (2009). Promoting Learning and Achievement Through Self-Assessment. *Theory into Practice*, *48*(1), 12–19. doi:10.1080/00405840802577544

Andrade, H. G., & Boulay, B. A. (2003). Role of Rubric-Referenced Self-Assessment in Learning to Write. *The Journal of Educational Research*, *97*(1), 21–30. doi:10.1080/00220670309596625

Arter, J., & McTighe, J. (2001). *Scoring Rubrics in the Classroom: Using Performance Criteria for Assessing and Improving Student Performance*. SAGE Publications.

Bolliger, D. U., & Halupa, C. (2018). Online student perceptions of engagement, transactional distance, and outcomes. *Distance Education*, *39*(3), 299–316. doi:10.1080/01587919.2018.1476845

Bolliger, D. U., & Martin, F. (2018). Instructor and student perceptions of online student engagement strategies. *Distance Education*, *39*(4), 568–583. doi:10.1080/01587919.2018.1520041

Brookhart, S. M. (2018). Appropriate Criteria: Key to Effective Rubrics. *Frontiers in Education*, *3*, 22. https://www.frontiersin.org/articles/10.3389/feduc.2018.00022. doi:10.3389/feduc.2018.00022

Brookhart, S. M., & Chen, F. (2015). The quality and effectiveness of descriptive rubrics. *Educational Review*, *67*(3), 343–368. doi:10.1080/00131911.2014.929565

Carroll, N., & Conboy, K. (2020). Normalising the "new normal": Changing tech-driven work practices under pandemic time pressure. *International Journal of Information Management*, *55*, 102186. doi:10.1016/j.ijinfomgt.2020.102186 PMID:32836643

Chaker, R., Bouchet, F., & Bachelet, R. (2022). How do online learning intentions lead to learning outcomes? The mediating effect of the autotelic dimension of flow in a MOOC. *Computers in Human Behavior*, *134*, 107306. doi:10.1016/j.chb.2022.107306

Chowdhury, F. (2018). Application of Rubrics in the Classroom: A Vital Tool for Improvement in Assessment, Feedback and Learning. *International Education Studies*, *12*(1), 61. doi:10.5539/ies.v12n1p61

Cullen, R., & Harris, M. (2009). Assessing learner-centredness through course syllabi. *Assessment & Evaluation in Higher Education*, *34*(1), 115–125. doi:10.1080/02602930801956018

Farrell, O., & Brunton, J. (2020). A balancing act: A window into online student engagement experiences. *International Journal of Educational Technology in Higher Education*, *17*(1), 25. doi:10.118641239-020-00199-x

Fraile, J., Panadero, E., & Pardo, R. (2017). Co-creating rubrics: The effects on self-regulated learning, self-efficacy and performance of establishing assessment criteria with students. *Studies in Educational Evaluation*, *53*, 69–76. doi:10.1016/j.stueduc.2017.03.003

Hughes, A. L., Michener, C., Mohamed, K., & Mcduff, N. (2019). Curriculum co-creation as a transformative strategy to address differential student outcomes: The example of Kingston University's Student Curriculum Consultant Programme. *Compass (Eltham)*, *12*(1), 1. Advance online publication. doi:10.21100/compass.v12i1.955

Jonsson, A. (2014). Rubrics as a Way of Providing Transparency in Assessment. *Assessment & Evaluation in Higher Education*, *39*(7), 840–852. doi:10.1080/02602938.2013.875117

Jönsson, A., & Panadero, E. (2017). The Use and Design of Rubrics to Support Assessment for Learning. In D. Carless, S. M. Bridges, C. K. Y. Chan, & R. Glofcheski (Eds.), *Scaling up Assessment for Learning in Higher Education* (pp. 99–111). Springer. doi:10.1007/978-981-10-3045-1_7

Jonsson, A., & Svingby, G. (2007). The Use of Scoring Rubrics: Reliability, Validity and Educational Consequences. *Educational Research Review*, *2*(2), 130–144. doi:10.1016/j.edurev.2007.05.002

Joseph, S., Rickett, C., Northcote, M., & Christian, B. J. (2020). 'Who are you to judge my writing?': Student collaboration in the co-construction of assessment rubrics. *New Writing*, *17*(1), 31–49. doi:10.1080/14790726.2019.1566368

Kahu, E. R., & Nelson, K. (2018). Student engagement in the educational interface: Understanding the mechanisms of student success. *Higher Education Research & Development*, *37*(1), 58–71. doi:10.1080/07294360.2017.1344197

Kilgour, A., Morton, J., Cloete, L., Dawson, S., & Northcote, M. (2022). *Rubric co-construction in medical and allied health education: Students' and teachers' perceptions* [Preprint]. In Review. doi:10.21203/rs.3.rs-1984776/v1

Kilgour, P., Northcote, M., Williams, A., & Kilgour, A. (2020). A plan for the co-construction and collaborative use of rubrics for student learning. *Assessment & Evaluation in Higher Education*, *45*(1), 140–153. doi:10.1080/02602938.2019.1614523

Lewis, E. (2021). Best Practices for Improving the Quality of the Online Course Design and Learners Experience. *The Journal of Continuing Higher Education*, *69*(1), 61–70. doi:10.1080/07377363.2020.1776558

Linnenbrink, E. A., & Pintrich, P. R. (2003). The Role of Self-Efficacy Beliefs In Student Engagement and Learning In the classroom. *Reading & Writing Quarterly*, *19*(2), 119–137. doi:10.1080/10573560308223

May, C. R., Albers, B., Bracher, M., Finch, T. L., Gilbert, A., Girling, M., Greenwood, K., MacFarlane, A., Mair, F. S., May, C. M., Murray, E., Potthoff, S., & Rapley, T. (2022). Translational framework for implementation evaluation and research: A normalisation process theory coding manual for qualitative research and instrument development. *Implementation Science : IS*, *17*(1), 19. doi:10.118613012-022-01191-x PMID:35193611

May, C. R., Cummings, A., Girling, M., Bracher, M., Mair, F. S., May, C. M., Murray, E., Myall, M., Rapley, T., & Finch, T. (2018). Using Normalization Process Theory in feasibility studies and process evaluations of complex healthcare interventions: A systematic review. *Implementation Science : IS*, *13*(1), 80. doi:10.118613012-018-0758-1 PMID:29879986

May, C. R., Mair, F., Finch, T., MacFarlane, A., Dowrick, C., Treweek, S., Rapley, T., Ballini, L., Ong, B. N., Rogers, A., Murray, E., Elwyn, G., Légaré, F., Gunn, J., & Montori, V. M. (2009). Development of a theory of implementation and integration: Normalization Process Theory. *Implementation Science : IS*, *4*(1), 29. doi:10.1186/1748-5908-4-29 PMID:19460163

McNair, T. B. (2016). Designing purposeful pathways for student achievement through transparency and problem-centered learning. *Peer Review : Emerging Trends and Key Debates in Undergraduate Education*, *18*(1–2), 3–6. https://go.gale.com/ps/i.do?p=AONE&sw=w&issn=15411389&v=2.1&it=r&id=GALE%7CA459505877&sid=googleScholar&linkaccess=abs

Mills, M. S. (2022). Promoting Inclusivity Through a Culturally Responsive Approach to Classroom Assessment Practices [Chapter]. Handbook of Research on Policies and Practices for Assessing Inclusive Teaching and Learning. IGI Global. doi:10.4018/978-1-7998-8579-5.ch018

Montgomery, K. (2002). Authentic Tasks and Rubrics: Going Beyond Traditional Assessments in College Teaching. *College Teaching*, *50*(1), 34–40. doi:10.1080/87567550209595870

Murdoch, D., English, A. R., Hintz, A., & Tyson, K. (2020). *Feeling Heard*: Inclusive Education, Transformative Learning, and Productive Struggle. *Educational Theory*, *70*(5), 653–679. doi:10.1111/edth.12449

Murray, E., Treweek, S., Pope, C., MacFarlane, A., Ballini, L., Dowrick, C., Finch, T., Kennedy, A., Mair, F., O'Donnell, C., Ong, B. N., Rapley, T., Rogers, A., & May, C. (2010). Normalisation process theory: A framework for developing, evaluating and implementing complex interventions. *BMC Medicine*, *8*(1), 63. doi:10.1186/1741-7015-8-63 PMID:20961442

Olesen, K. B., Christensen, M. K., & O'Neill, L. D. (2020). What do we mean by "transferable skills"? A literature review of how the concept is conceptualized in undergraduate health sciences education. *Higher Education*. *Skills and Work-Based Learning*, *11*(3), 616–634. doi:10.1108/HESWBL-01-2020-0012

Panadero, E., & Jonsson, A. (2013). The use of scoring rubrics for formative assessment purposes revisited: A review. *Educational Research Review*, *9*, 129–144. doi:10.1016/j.edurev.2013.01.002

Panadero, E., Jonsson, A., & Botella, J. (2017). Effects of self-assessment on self-regulated learning and self-efficacy: Four meta-analyses. *Educational Research Review*, *22*, 74–98. doi:10.1016/j.edurev.2017.08.004

Peacock, S., Cowan, J., Irvine, L., & Williams, J. (2020). An Exploration into the Importance of a Sense of Belonging for Online Learners. *International Review of Research in Open and Distance Learning*, *21*(2), 18–35. https://eric.ed.gov/?id=EJ1250669. doi:10.19173/irrodl.v20i5.4539

Ragupathi, K., & Lee, A. (2020). Beyond Fairness and Consistency in Grading: The Role of Rubrics in Higher Education. In C. S. Sanger & N. W. Gleason (Eds.), *Diversity and Inclusion in Global Higher Education: Lessons from Across Asia* (pp. 73–95). Springer. doi:10.1007/978-981-15-1628-3_3

Reddy, Y. M., & Andrade, H. (2010). A review of rubric use in higher education. *Assessment & Evaluation in Higher Education*, *35*(4), 435–448. doi:10.1080/02602930902862859

Reynolds-Keefer, L. (2010). Rubric-referenced assessment in teacher preparation: An opportunity to learn by using. *Practical Assessment, Research & Evaluation*, *15*(1), 8. doi:10.7275/PSK5-MF68

Rezaei, A. R., & Lovorn, M. (2010). Reliability and validity of rubrics for assessment through writing. *Assessing Writing*, *15*(1), 18–39. doi:10.1016/j.asw.2010.01.003

Sadler, D. R. (2005). Interpretations of criteria-based assessment and grading in higher education. *Assessment & Evaluation in Higher Education*, *30*(2), 175–194. doi:10.1080/02602930420000264262

Sadler, D. R. (2011). Academic freedom, achievement standards and professional identity. *Quality in Higher Education*, *17*(1), 85–100. doi:10.1080/13538322.2011.554639

Sadler, D. R. (2013). Assuring academic achievement standards: From moderation to calibration. *Assessment in Education: Principles, Policy & Practice*, *20*(1), 5–19. doi:10.1080/0969594X.2012.714742

Skillings, M. J., & Ferrell, R. (2000). Student-Generated Rubrics: Bringing Students into the Assessment Process. *The Reading Teacher*, *53*(6), 452–455.

Snelling, C. A., Loveys, B. R., Karanicolas, S., Schofield, N. J., Carlson-Jones, W., Weissgerber, J., Edmonds, R., & Ngu, J. (2019). Partnership through co-creation: Lessons learnt at the University of Adelaide. *International Journal for Students as Partners*, *3*(2), 2. doi:10.15173/ijsap.v3i2.3799

Stevens, D. D., & Levi, A. J. (2004). Introduction to Rubrics: An Assessment Tool to Save Grading Time, Convey Effective Feedback and Promote Student Learning. In *Stylus Publishing, LLC*. Stylus Publishing, LLC.

Torrance, H. (2007). Assessment as learning? How the use of explicit learning objectives, assessment criteria and feedback in post-secondary education and training can come to dominate learning. *Assessment in Education: Principles, Policy & Practice*, *14*(3), 281–294. doi:10.1080/09695940701591867

Ulker, V. (2017). The Design and Use of Speaking Assessment Rubrics. *Journal of Education and Practice*.

Voltmer, J.-B., Reich-Stiebert, N., Raimann, J., & Stürmer, S. (2022). The role of multi-attributional student diversity in computer-supported collaborative learning. *The Internet and Higher Education*, *55*, 100868. doi:10.1016/j.iheduc.2022.100868

Watty, K., Freeman, M., Howieson, B., Hancock, P., O'Connell, B., de Lange, P., & Abraham, A. (2014). Social moderation, assessment and assuring standards for accounting graduates. *Assessment & Evaluation in Higher Education*, *39*(4), 461–478. doi:10.1080/02602938.2013.848336

Wiggins, G. (1990). The Case for Authentic Assessment. *Practical Assessment, Research & Evaluation*, *2*(1), 2. doi:10.7275/FFB1-MM19

Wiliam, D. (2000). *Integrating Summative And Formative Functions Of Assessment*.

Wood, P. (2017). Overcoming the problem of embedding change in educational organizations: A perspective from Normalization Process Theory. *Management in Education*, *31*(1), 33–38. doi:10.1177/0892020616685286

Zimmerman, B. J. (2000). Attaining Self-Regulation: A Social Cognitive Perspective. In M. Boekaerts, P. R. Pintrich, & M. Zeidner (Eds.), *Handbook of Self-Regulation* (pp. 13–39). Academic Press., doi:10.1016/B978-012109890-2/50031-7

KEY TERMS AND DEFINITIONS

Co-Creation: Collaboration among different stakeholders to produce an outcome that encompasses the ideas, knowledge, interests, and perspectives of all participating parties.

Deep and Reflective Learning: A process in which the student has ownership of their learning by actively reflecting on their previous performance, developing an understanding of given requirements as

well as their current level of achievement, and using feedback/feedforward to direct their future learning and performance.

The IoPPN's Co-Production (Co-Pro) Student Panel: A group that represent the IoPPN's online students from the Master programmes 'Psychology and Neuroscience of Mental Health' and 'Applied Neuroscience' in regular meetings with staff to discuss and develop new academic and extra-curricular initiatives.

Normalisation Process Theory (NPT): A conceptual framework guiding the design and the integration of complex interventions into routine and already established processes.

Online Education: Higher education programmes that are delivered 100% online in which the instructors and students interact within a virtual learning environment.

Rubric: A scoring tool composed of a set of criteria/descriptors and an accompanying scale illustrative of differing levels of student performance.

Stakeholder: Interested parties who can affect or be affected by the outcome of co-creation (students, educators, IoPPN).

APPENDIX 1

Table 4. CO-PRO panel: definition and recruitment questions

The CO-PRO panel exists to involve students in course design, material development and other related tasks as relevant. This reflects KCL's student-centred approach. The panel will work in an open, transparent and inclusive way, respecting the diversity of perspectives among panellists and the wider student population. **We would like to encourage students from diverse backgrounds to apply to make our panel representative.**	
My KCL email address is	
I am a student on the following Course:	
I am a student on the following carousel:	
We are keen that the CO-PRO panel is representative of our diverse community of online students at the IoPPN. Can you tell us a little bit about yourself, and why you are interested in this role? (300 words max)	
What experience do you have in representing others (this does not have to be in Higher Education)? If none, do not worry - please let us know what skills you have that would make you a good candidate.	
The deadline for this survey is the 5th of September. We will notify you shortly thereafter whether or not you have been selected to be on the panel.	

Chapter 9

Facilitating Programme–Level Assessment Working Teams to Develop Shared Rubrics Across a UG and PG Programme Portfolio in Business Education

Roisin Donnelly

https://orcid.org/0000-0001-9053-6375
Technological University Dublin, Ireland

Colin Hughes

Technological University Dublin, Ireland

ABSTRACT

This chapter is a reflective study reporting on a College-wide common rubrics initiative in a Technological University (TU) in Ireland. Assessment and feedback are enduring issues for the higher education sector both in Ireland (as well as internationally). By addressing these priorities, we are focusing on the connected areas of marking practices and feedback processes in a College of Business. The chapter highlights the collaborative nature of an initiative on programmatic assessment design, its breadth of scope, and the high levels of support provided to staff and students through the design process. In particular, rubrics are the main focus of the chapter with an overview of Programme Learning Outcome (PLO) mapping provided as part of the context. Four interdisciplinary rubric working teams were formed across the College of Business to develop common rubrics in the areas of reflective practice, critical thinking, individual/group presentation skills and industry consultancy projects. This programme-based study differs from other previous work as it involves working on a consistent basis with the challenges of bringing cultures, practices and understandings of disciplinary teams together in a technological uni-

versity context. Findings from our collaborative common rubric working teams showed the importance of avoiding designing overly complex rubrics and of focusing on the student's work providing evidence of meeting, exceeding, or falling short of the quality being looked for. It also highlighted the importance of rubric literacy as we were using qualitative language in all the target statements and advised staff to avoid subjective language in favour of explicit guidance. We proposed the need for a common language on 'Teamwork' and on proceeding with this as a College-wide approachd a similar structure or focus, using the same marking rubric was recommended. We advised staff to work as a team to master the rubrics, and to ensure that their students had m. Future work can explore the use of AI tools which can automate the feedback process and provide lecturers with customised rubrics based on their specifications.

INTRODUCTION

Although working collaboratively on rubric development is well-documented in the literature (Allen & Knight, 2009; Morton *et al*., 2021), the seminal work of Sadler (2009) and Andrade (2005) was useful in turn for our specific context. The former for contributing to our understanding of how to design assessments to enable learners to demonstrate sophisticated cognitive abilities, integration of knowledge, complex problem solving, critical reasoning, original thinking, and innovation. The latter for using rubrics with undergraduate and postgraduate students.

In the context of this collaborative rubric initiative across a College of Business, there has been a clear intention to address grading, based on a realisation that the grade alone does not provide sufficiently informative feedback to students. Jutras (2023) argues that the question of grading schemes is a recurring topic in academic discussions where the key word is evaluation, in this instance of assignments and exams. By taking a college-wide approach to designing and using rubrics in common areas that deepen student thinking and learning, the aim of this collaborative work was to:

- help students improve their work by making assessment expectations explicit and aiding the feedback process; Panadero & Jonsson (2013) have previously reported that rubrics help learners clarify expectations, reduce anxiety, and improve self-efficacy.
- support student learning in an environment where independence and the ability to self-regulate is important; a key part of the context for the focus on rubrics is the College's signature pedagogy on student-centred learning, encouraging students to take ownership of their learning and be more self-directed. As Sadler (2009) explained, the aim is for learners to become better able to engage in self-monitoring the development of their own works.
- aid programme teams in achieving higher levels of consistency when assessing student work; Jonsson & Panadero (2017) discussed how rubrics can help with reliability of grading when multiple markers are involved, and thus help with moderation of markers and quality assurance processes.
- improve efficiency for staff; it is known that rubrics are attractive to busy academics as they can save considerable time, especially as they reduce the number of queries from students - which was previously discussed by Honeychurch (2015) in a Business School context.

The context for this project is what was previously the College of Business and is now the Faculty of Business in Ireland's first Technological University **-** TU Dublin **-** which came into existence in January

2019, merging three existing higher education institutions. Technological universities offer programmes that are vocationally and professionally oriented. The Faculty of Business is one of Ireland's largest Business Schools in terms of full-time and part-time student enrolment and has been in existence for over thirty years. The Faculty offers undergraduate, postgradauate and executive education programmes to approximately 7,000 students across a range of Business disciplines. The five Schools at the centre of this project were the Graduate Business School, Schools of Management, Marketing, Retail, and Accounting & Finance.

As can be imagined for this Business Faculty with its extensive history, moving into a new Technological University context has meant that there is significant change happening to the existing institutional structure and fabric, and such system-wide restructuring can have a profound impact for students and staff. Against this backdrop of institutional change, the Faculty of Business is continuing to hone its professionally oriented national and international programme provision for students and reflect on how it designs programmes to meet society and industry's needs for the future.

BACKGROUND

There were two parallel and interlocking assessment and feedback collaborative projects taking place in the College:

- A Programme Learning Outcome (PLO) mapping initiative
- College Level PLO Rubric project

The use of rubrics had been increasing steadily within the College, with faculty members becoming more comfortable with their development and implementation and reporting several benefits for both students and staff. In an effort to extend the benefits of rubric usage across the College, the Head of Learning Development established a number of collaborative rubric working teams. A decision was made to focus on Programme Learning Outcomes (PLOs) and this was influenced by several factors.

Firstly, the College was grappling with an issue of over assessment for a number of years. Previous studies on over-assessment provide useful background for our work. A study by Jessop *et al.* (2014), analysing 23 degree programmes in eight universities, offers an approach (TESTA) for helping teachers to redesign assessment regimes. A study by Jessop & Tomas (2016) on 73 programmes in 14 UK universities reported that students typically encounter eight times as much summative as formative assessment, a dozen different types of assessment with more than three quarters by coursework. They argue that high varieties of assessment are probable contributors to student confusion about goals and standards. These findings build on an earlier study by Gibbs & Dunbar-Goddet (2009) who reported that programmes were found to have either a high volume of summative assessment or a high volume of formative-only assessment, but never both at the same time. Harland & Wald (2021) placed the rationale for increased assessment levels on semesterisation and a module structure emerging in the 1990s in their New Zealand HE context, arguing that assessment levels have remained high and unchanging since. This mirrors the modular structure of the Irish HE sector which began in the mid-1990s and had been introduced to allow students the option of spreading their assessment and examination workload across the academic year, thereby reducing pressure at the end of the year.

In our College, at the majority of programme examination boards, external examiners highlighted areas to address to alleviate the workload, for either upcoming or future marking, as well as making marking (or the results thereof) more meaningful. Students have continuously asked for lower quantity of continuous assessments and more in-depth feedback to show their progress across their programme of study. From the academic staff' perspective, the issue was that lecturers were spending large proportions of their time marking/grading that is created from asking students to demonstrate their skills often, at length, and in detail in their programme of study. Every discussion post, short-response essay, exam, annotated bibliography, individual/group work, and capstone project was weighed, ranked, and assigned a numerical grade. While this rigorous approach to assessment challenged students and ensured the achievement of learning outcomes, the College needed to consider ways of reducing the workload for both students and lecturers while maintaining rigour and quality.

Secondly, the College was preparing to seek accreditation from the Association to Advance Collegiate Schools of Business (AACSB). AACSB provides quality assurance, business education intelligence, and learning and development services to over 1,850 member organizations and more than 950 accredited business schools worldwide (AACSB, 2022) and is one of three major Business Schools accreditations, along with AMBA (Association of MBAs who have traditionally focused on accrediting MBA programmes at Business Schools but have recently extended their remit with BGA accreditation) and EQUIS (EFMD Quality Improvement System which is the leading international system of quality assessment, improvement and accreditation of Business Schools).

Beginning the PLO Mapping Process

Along with mapping PLOs, programme teams were required to map assessments, with a dual purpose of AOL mapping and of addressing the issue of over-assessment. Detailed PLO mapping and discussion proved to be immediately helpful in creating a more efficient assessment strategy. For instance, in one particular programme, several modules were assessing reflective writing skills, which led to unnecessary workload for students and faculty. Using the IRMA tool provided lecturers with greater clarity across the programme, with the result that one lecturer was content to stop assessing reflective writing skills, safe in the knowledge that it was being assessed elsewhere, and could focus instead on simply reinforcing what had already been introduced to students. In another module, where students were required to present on a team project, the lecturer realized that they no longer had to assess presentation skills, as they were adequately assessed elsewhere, but rather they could focus on assessing other module learning outcomes.

As part of AACSB preparation, the College was seeking to further enhance its approach to Assurance of Learning (AOL), a key requirement of AACSB accreditation (Standard 5) (AACSB, 2020). AOL processes measure what a student has learned upon completion of their programme (against stated learning competencies), identify areas where student learning is deficient, and outline changes to the curriculum and learning experience to ensure that learning competencies are met (Tarnoff, 2023; AACSB, 2020

A key principle of AOL is the measurement of PLO achievement and a process by which the programme team is seen to 'close the loop', designing for PLO achievement, reflecting on student performance against PLOs, making the necessary changes to the programme before measuring again. Schools are expected to measure twice in a five-year cycle, with improvements made between two measurement cycles to improve the curriculum. As such, Schools typically "close the loop" at least once in a five-year cycle (AACSB, 2019).

Curriculum mapping plays a central role in the AOL process, allowing programme teams to audit the achievement of learning outcomes or objectives, while simultaneously providing students with information which allows them to play a more active role in their learning (Biggs, 1999; Harden, 2001). In writing about quality control in curriculum development, English (1978) proposed the three categories of declared, delivered and learned curriculum. This seminal work influenced a significant volume of further research, with the field of curriculum mapping expanding significantly. For instance, Cuban (1993) proffered three different labels of intended, taught and learned curriculum while also adding a fourth category, which they labelled tested (Cuban, 1993), while Jacobs (1997) built on English's (1978) work in creating a seven-stage model of curriculum mapping.

While much of the earlier work on curriculum mapping was focused at the school level, research has since expanded within the Higher Education context (Tariq *et al.*, 2004; Sumsion & Goodfellow, 2004; Robley, Whittle & Murdoch-Eaton, 2005a, 2005b; Spencer *et al.*, 2012; Okojie *et al.,* 2022). Joyner (2016) argues that curriculum mapping and assessment will increasingly be required as Universities seek quantitative data to monitor college and departmental performance. However, it is also a crucial consideration for Business Schools committed to obtaining international accreditations such as AACSB. As discussed, AOL requires business Schools to demonstrate how Programme Learning Goals/Outcomes are achieved. As the planned curriculum does not always align with the operational curriculum (Hale, 2008) AACSB like to see clear evidence of curriculum mapping and assessment (along with enhancement).

In order to help them with this mapping process, Business Schools generally use a 3 or 4 step mapping process. Hale (2008) identified three stages of learning, introductory (where students are first exposed to the material or topic), developing (where students strengthen their understanding) and mastery (where the students have a full understanding of the topic). A variety of this model, referred to as IRMA is commonly used by AACSB accredited Business Schools, a sample of which are shown in Table 1. Based on the understanding that it takes more than a single course/module to develop certain skills (Tarnoff, 2023), the IRMA model helps Schools to map (across a programme) where a PLO is: introduced, reinforced or where students have an opportunity to practice, and where the student has had sufficient opportunities to practice and should now be in a position to demonstrate mastery. Some Schools only use I, R and M, whereas others use A to specify where the PLO is assessed.

Table 1. University/Business school usage of IRM/A model

University	Categories			
Berkeley	Introduced	Reinforced/Practiced	Mastered	Assessed
University of Hawaii	Introduce	Reinforce	Master	Assessment
University of Tennessee at Chattanooga	Introduced	Reinforced/Practiced	Mastered & Assessed	
University at Buffalo	Introduced	Reinforced and/or Practiced	Mastered	
Western Kentucky University	Introduced	Reinforced	Mastered	Assessed

While many other alternatives to IRMA exist, including T/D/M/A (Training, Development, Monitoring, Assessment (Gibbs *et al.,* 2004) and I/E/U/A (Introduced/Emphasized/Utilized/Comprehensive

Assessment (Stassen *et al.*, 2011) to name but two, we decided to use IRMA as it had been previously introduced to colleagues and was used widely by AACSB accredited Schools. A key area for agreement centred on the concept and practice of 'Mastery' across programmes and how to assess for Mastery. As such, conversations with our programme teams took place on:

What does mastery look like in each discipline?
Are we assessing for content and skills mastery only in final year?
What happens if a discipline skill remains at introductory level e.g. is not measured for mastery?

ISSUES, CONTROVERSIES, PROBLEMS

This section discusses Programmatic Assessment and Feedback Design, with a focus on rubrics in Business Education. While the broad rationale for our work was a College-wide programme level assessment design and feedback strategy, the specific contribution of rubrics in this chapter are rubric literacy, and the plans about how this will be co-developed with staff and students, which are explored in the next section.

Introducing Rubrics

There were pockets of excellence with regard to rubric usage in the College (such as a financial reporting rubric which was detailed and useful for feedback purposes), and other rubrics had been developed and implemented by programme teams to help with the development of skills such as reflective writing and presentation skills. However, when we initially gathered together all existing rubrics for review it became apparent that some were not designed to provide adequate feedback to students. These rubrics were revisited by some teams as they set out on their PLO mapping journey and while the IRMA tool provided assurances to lecturers that key skills/ competencies, were being developed in a planned fashion, rubrics provided assurances around the quality and timeliness of feedback and the potential to further encourage self-directed learning.

While some rubrics were already in place, they: (a) had not always benefitted from a deep peer review at the design phase, (2) were by no means used consistently across programmes and (3) they had not been developed in line with the new PLOs. Therefore, the rubrics project was simultaneously launched to address these issues and ensure that each programme team could consistently use rubrics to help students to achieve PLOs. In differentiating between formative and summative assessment, and conscious of the need to allow students time to reinforce or practice, it was hoped that rubrics would be used at each stage of the learning journey, and not just when assessing for mastery. For instance, a presentation skills rubric could be used to provide feedback to a student consistently across their learning journey from the time of introduction, to reinforcement/practice, to mastery. It was also hoped that the project would create greater awareness of rubric use, provide opportunities for peer learning and discussion and also give support to academic staff looking to enhance their use of rubrics at a module level.

Core within this collaborative activity, multi-disciplinary teams were encouraged to reimagine and redesign rubrics for their programmes. Early on in the process, they were presented with a body of research and evidence on the pros and cons of rubrics, and lessons learnt from the literature. Reddy & Andrade (2010) critically reviewed the empirical research on the use of rubrics at post-secondary level and found that studies of the validity of rubrics have shown that clarity and appropriateness of language

is a central concern. Rubric literacy is a key consideration at the beginning of a process with staff and students. Dawson (2017) highlights that 'Rubric' is a term with a variety of meanings, and that over time it has come to represent divergent practices. This is significant to acknowledge as rubrics have been evaluated, mandated, embraced and resisted based on often imprecise and inconsistent understandings of the term. Brookhart (2018) also highlighted confusion that can exist when some studies called their assessment tool a "rubric" when in fact it was a rating scale, and in some rubrics, performance level descriptions used rating-scale language or counted occurrences of elements instead of describing quality.

However, in their more recent critical review of the arguments against the use of rubrics, Panadero & Jonsson (2020) concluded that rubrics seem to have more benefits than drawbacks, especially when used formatively. A useful study for our Business Education context was from Garrett *et al.* (2012) who found in a small Business School in the U.S. that methods driven and developed by faculty members can lead to a stronger 'bottom-up' approach to quality assurance, and the major lesson learned from their study is that there is no one single "right" way to do assessment.

PROCESS

Our primary rationale was ensuring that we were enabling students to achieve PLOs by providing clarity of expectations and clear feedback on performance. Such focus also helped with over-assessment as lecturers were assured that students were developing competencies relating to specific PLOs in earlier modules and could therefore refocus on other module level outcomes and remove certain MLOs, content and assessment.

How peer-led teams worked collaboratively to co-construct and review common rubrics

To initiate our new rubric work, we invited an international expert in the field of feedback to speak with all academic staff in the College on a number of key areas for understanding before the process got underway in the rubric working teams: the theoretical underpinnings of rubrics for assessment and feedback; the anatomy of a rubric; holistic v analytic rubrics; addressing consistent rubric design and use across UG degree programmes (what is the best approach to bring programme teams together to work on this; what design stages are needed); what does a year 1 rubric vs a year 4 rubric holistic vs analytic look like in a programme; avoiding over-assessment for students – how to plan for this as a programme team; how do rubrics differ for Business students; how can they be used to show advancement in Business practice (what does this involve? What does it look like?); and finally to share a number of Good Practice examples of rubrics in critical thinking and reflective practice.

The shared rubric design process (Figure 1) was aimed at both UG and PG degree-programme teams, who worked in parallel during the process and shared good practice as they re-thought their programme assessment strategies and undertook their collaborative design work.

Figure 1. Stages of the collaborative rubric working process
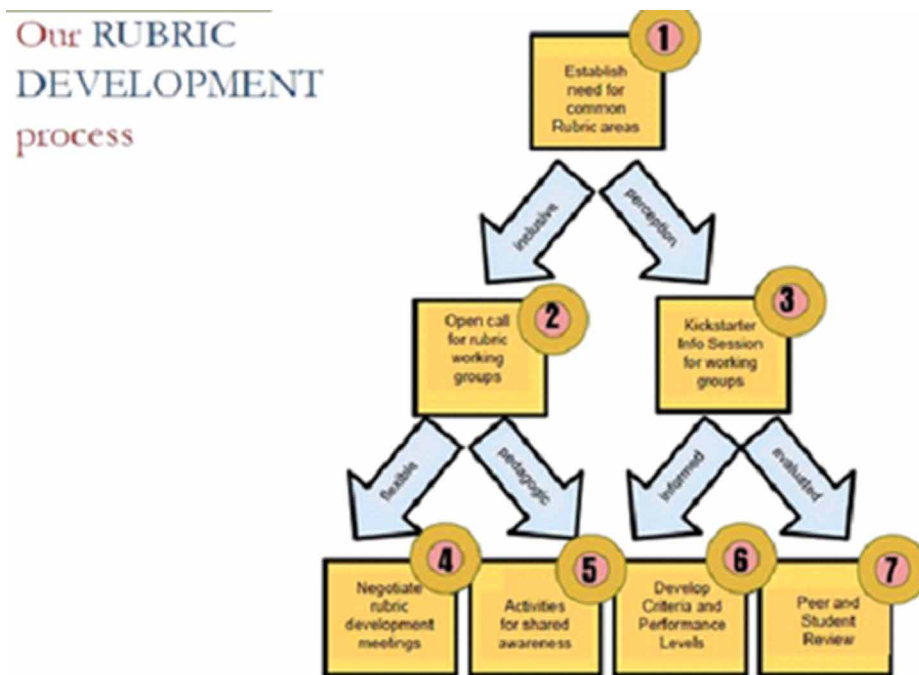
The specific rubric projects which is disseminated throughout the chapter are drawn from initiatives in community-based work, developments on work placements, digital assessment practices, as well as a selection of activities to develop digital literacy, critical thinking, teamwork and reflective practice in advance of final year capstone projects. These emerged from dual sources – over recent years, across a number of programmes in the School, external examiners in their quality assurance reports identified a need for reduction in assessments and improved feedback processes. At the same time, student feedback and staff experiences reported similar findings (Figure 2). This figure shows the rubric working teams at the centre of our process. Surrounding those, the process involved key considerations such as beginning with a challenging assignment, sustained inquiry, authenticity, student voice and choice, reflection, critique and revision, before ending with a public product.

Figure 2. Considerations informing the work of the rubric teams

A number of support initiatives were introduced to the collaborative rubric process. It was recognised that important work had already taken place in niche areas of the College but had not been widely disseminated. A first step was to address this, and a reflective practice 'swap-shop' was held where all programme teams were asked to identify examples of good practice in their Annual Quality Assurance report and representatives were invited to present to each other. The purpose of the Swap-Shop was to enable staff in the College to share their experience of designing and delivering good reflective practice opportunities for their students taking place on their modules/programmes. It provided an opportunity for colleagues to discuss, exchange, source and share ideas and good practices related to reflective practice in Business Education and consider how we could enhance these whilst supporting the development of our students.

Social construction of rubrics was important in our approach in collaborative working in teams, was the underpinning theoretical rationale, and formed the basis for activities in which lecturers and students engage to share common motives and work towards a common goal. Collaboration between the teams of academic staff in this initiative was a powerful professional development activity that helped the staff not only improve their assessment and feedback strategies in different ways but also learn new ideas to try in their practice. The individual lecturers were actively constructing their own knowledge on rubric design and development by way of experiences and interactions with others in their rubric teams. As Ragupathi & Lee (2020) report, rubrics provide educators with a greater understanding of their own teaching practice and encourage them to become reflective practitioners.

We want to now highlight the process we undertook and the lessons learnt from this experience in the form of a series of implications for other practitioners in Business disciplines.

In taking its initial steps on a journey towards AACSB accreditation, the College leadership team developed a set of Programme Learning Outcomes (PLOs) for Level 8 (Bachelor) and Level 9 (Master) programmes (see Tables 2 and 3). This work was informed by a peer review of AACSB accredited

schools, desk-based research on required graduate attributes and competencies, engagement with employers and colleagues. A set of seven PLOs was agreed for Bachelor programmes and six PLOs for Master programmes. Programme teams were then asked to adopt or adapt these PLOs for each individual programme. While the PLOs related to a number of different competencies, several were common and this required the rubric working teams to discuss the requirements of students at different levels, itself a very worthwhile exercise as it forced dep discussion about Faculty's expectations of students at different levels and what mastery looked like at each level.

As AOL outlines a process for mapping programme level learning competencies, initially this is where the College focused its energy, producing a set of common Programme Learning Outcomes (the College uses the language of Outcomes in alignment with Irish Quality Assurance standards) for both UG and PG programmes. These were distributed to Programme Chairs, along with a learning outcomes mapping template a Mapping Process Support Guide which Programme Chairs could use with their programme teams to reflect on and stimulate discussion about Programme Learning Outcomes (PLO) mapping. While the PLOs developed at College level were designed to be common across programmes at each level, Programme Teams were given licence to adapt and nuance the PLOs for their programmes.

Table 2. Undergraduate programme learning outcomes

Undergraduate Learning Goal/Outcome
Communication: Students will be able to research, organise, present, deliver and write an effective document in a professional manner.
Critical Thinking/ Problem Solving/ Entrepreneurial Decision Making: The student will be able to identify and analyse problems and devise appropriate solutions.
Teamwork: The student will be able to participate effectively in teams.
Using Information Technology: The student will be able to use existing technology effectively and have the skills necessary to adapt and apply new technologies.
Ethics and Corporate Social Responsibility: The student will be able to identify aspects of ethical dilemmas from multiple stakeholders' perspectives and offer viable alternative solutions.
Cross-Disciplinary Knowledge: The student will demonstrate the general/ core supporting knowledge relevant to the field of study (e.g. business modules such as foundation level marketing, accounting, management, economics, etc.)
Discipline-Specific Knowledge: The student will demonstrate the knowledge required to obtain an entry-level position in the discipline (i.e. profession)

Table 3. Postgraduate programme learning outcomes

Postgraduate Learning Goal/Outcome
Communication: *Programmes develop each graduate to be a persuasive communicator and negotiator.* Students will be able to use a range of communications strategies to reach agreement with others about appropriate responses to complex and unfamiliar problems within one or more fields of business practice.
Business Analysis & Problem Solving: *Programmes develop each graduate to be a capable business analyst & strategic problem-solver.* Students will be able to apply a range of quantitative & qualitative research skills to identify & diagnose complex, unfamiliar problems & to use the evidence & findings generated to formulate strategically appropriate solutions within one or more fields of business practice.
Critical Thinking: *Programmes develop each graduate to be an autonomous & constructive critical thinker.* Students will be able to question, assess & respond independently & creatively to assumptions, propositions & debates within one or more fields of business practice.
Teamworking: *Programmes develop each graduate to be a capable team leader in work-related contexts.* Students will be able to influence others to work collaboratively to address complex and unfamiliar problems within one or more fields of business practice.
Business Knowledge: *Programmes develop each graduate to be a knowledgeable business practitioner.* Students will be able to demonstrate an integrated understanding of key concepts, techniques & trends in one or more fields of business practice & the challenges & opportunities involved in applying this knowledge in diverse contexts.
Ethics & Social Responsibility: *Programmes develop each graduate to be an ethically- and socially- responsible professional.* Students will be able demonstrate ethical & social awareness & responsibility in personal decision-making & behaviour within one or more fields of business practice.

A group of staff had previously commenced work on a broader teamwork project, which encompassed an analysis of team-working guides, rubrics and current approaches to peer assessment, including the allocation of marks to student members in the team. Therefore, teamwork was not prioritised within this new rubrics initiative. Furthermore, it was felt that the Discipline/Business Knowledge outcomes would have to be discussed and developed under our new discipline structures and that this would be more of a medium-term project. However, there was broad agreement that rubrics could be created for a range of PLOs that would benefit every programme in the College.

A call went out to staff across the Schools in the College inviting participation in a series of working teams for developing college-wide rubrics in a number of identified areas:

Reflective Practice, Critical Thinking, Presentation Skills, Consultancy Projects (shown in Figure 3).

Seeking out collaborators was a key feature of this work; collaboration promotes engagement and teamwork in the workplace across Schools which can lead to increased faculty productivity. Logan *et al.* (2011) argue that natural teams can be formed among persons of like interests, passion, and common goals. Members encourage and motivate each other; relationships involve high levels of engagement, innovation, creativity, and pursuit of a greater good. This is key to collaboration. Staff self-elected to work together and we agreed a lead for each working team. However, no matter who took the lead, the work was always circulated to the others multiple times for input.

Figure 3. Shared rubrics developed across business disciplines
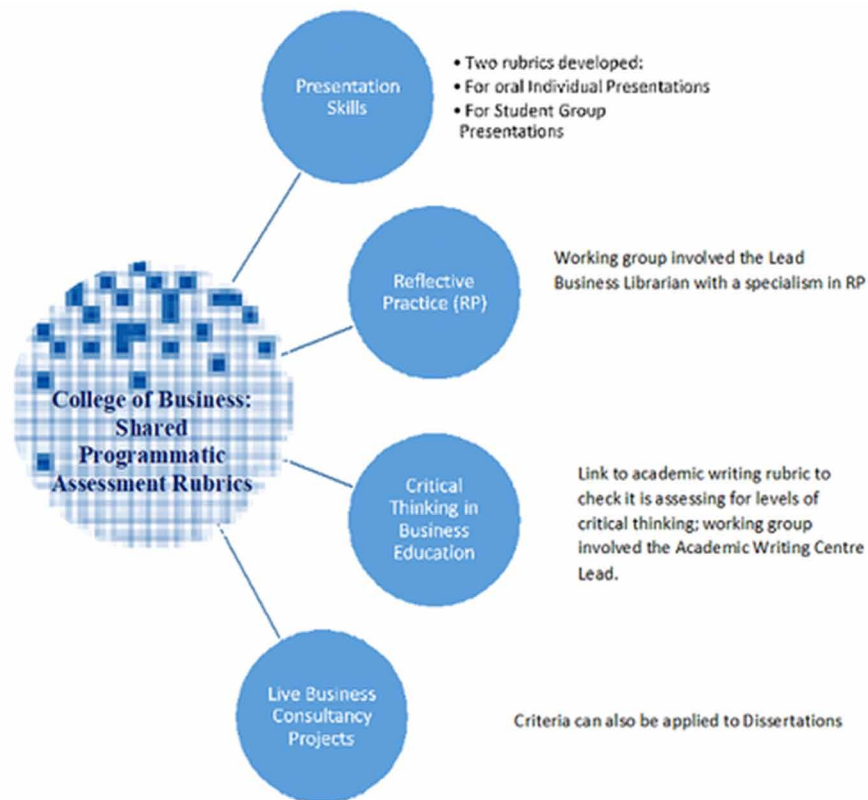
These rubric teams were formed and invited to a rubrics kick-starter session where the process for rubric development was discussed and agreed. A rubric resource pack to support the process was made available on the TLA Portal for staff. Over a number of weeks, the rubric teams were advised to meet to explore the repository of rubric examples made available to them in each of the identified areas, and to discuss issues arising in relation to the development of a common rubric for UG programmes across the College. At this initial kick-starter session, it was discussed and agreed what our purpose was for using common college-wide rubrics: that we needed a consistent approach to feedback (based on student feedback, external examiner reports and AACSB alignment) in order to make learning and assessment expectations explicit for students and to foster Feedback Literacy. Price *et al.* (2012) suggest that feedback can only be effective if the learner understands the feedback and is willing to act on it.

A collaborative activity in this session was very useful with each of the rubric working teams to build a shared understanding of the task among the team. As such, each team was asked to consider the definitions of their rubric areas and agree the purpose of their common rubric for both students and staff, and before, during and after the assessment process. Figure 4 shows the activity for the Critical Thinking (CT) rubric team.

Figure 4. Critical thinking activity underpinning the work of the CT rubric team

A rubric review session was then held bringing all the teams back together, with each team invited to share both their working process to date and their draft rubric. To support the process, we developed a simple resource-sharing rubric repository in the TLA Portal in the VLE (Figure 5). Resources, including rubric exemplars related to the topics were shared within this repository and colleagues were asked to supplement the repository with any useful resources that they discovered. We actively sought input from each other's professional domain.

Ultimately an eResource was developed to be used alongside each common rubric to support staff in having valuable conversations with students about their progress across programmes, and was made available from the Portal also. The eResource formed part of an 'ideas bank' of how best to use the rubrics e.g. to have students self-evaluate a presentation or evaluate one from their peers. This 'opens the

door' to critical thinking and construction of meaning and quality by students that is important at UG and PG levels. Participants in the working teams role was as a peer reviewer of the eResource. Simply handing out a rubric cannot be expected to have an impact on student work (Andrade & Valtcheva, 2009); students must be taught to actively use a rubric for self and peer assessment, and revision, in order to reap the benefits (Reddy & Andrade, 2010). In the wider context of student feedback literacy, Carless & Boud (2016, 1316) argue "students respond to feedback in various ways within specific disciplines, curricula and contextual settings…and in relation to their previous experiences and their own personal characteristics".

Figure 5. TLA portal to support staff in rubric development
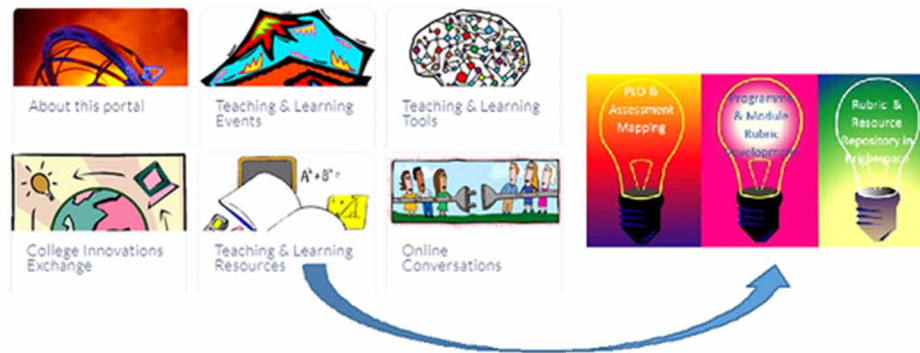
Early on, we realised the need for provision of clarity with all staff involved in the project at each stage of the project and we ensured that communications happened on a regular basis to convey the extent of the work that was taking place. As they were developed, the draft rubrics were uploaded to the Brightspace virtual learning environment in a 'College Innovation Exchange' portal, as an opportunity for all staff to provide online feedback on the draft rubrics. A second review meeting was scheduled with the rubric working teams discussing all the draft rubrics in relation to key considerations in rubric design (with input from our Learning, Teaching and Assessment team in the University). A rubrics information session was held in the Programme Chairs Forum with staff invited to show peers how to use the developed rubrics with intelligent agents in Brightspace and agreeing the nature of the accompanying resource explaining how best to use each rubric. A cross-university peer review group was set up for the rubrics, involving other Heads of Learning Development, colleagues from the central Learning and Teaching Centre, and lecturers from other disciplines with experience of rubric design – this was done as a sanity check for rubric structure and content.

To date, four of the five common rubrics have been created (Reflective Practice; Individual & Group Presentation Skills; The Consultancy Project). The Critical Thinking rubric has been developed in draft and needs to undergo peer review.

The literature emphases student learning as part of the development of rubrics (Reddy & Andrade, 2010; Huball & Burt, 2004; Olsen & Krysiak, 2021). For our context the ways in which student involvement in the process will get underway is with a plan to pilot the rubrics with students on particular programmes with the purpose of getting their feedback on identifying what quality should look like in

each of the areas, the language of the rubric (rubric literacy), and how it engages them with the feedback given. Training sessions will be put in place for the students to use the rubrics. Following this, staff will be invited to a lab session on inputting the final college rubrics to Brightspace, and get guidance on how each lecturer can best use them in their programme.

Challenges

As with any undertaking, there were challenges encountered along the way. They included finding the time to work on our rubrics, deciding when something was "good enough" for sharing with staff (even drafts!), maintaining commitment to the rubric project to ensure we kept moving forward, and deciding the performance levels in the rubric to work on next. With constrained resources, we were as Gabriel (2017) stated, asking staff to do more with less time, money, and personnel – and were aware that this can contribute to staff feeling overwhelmed. A study by Francis (2018) was useful for us as a College working together to create common rubrics to examine the link between rubrics and performance through the lens of student engagement. Findings from that study show providing a rubric does not necessarily lift student performance, whereas higher grades are evident when students engage with discussing the rubric and grades were ever higher when students engaged with the rubric discussion plus additional resources.

Rahmat (2017) outlined some challenges when designing a rubric that could empower students to self-assess their learning - addressing teachers' misconceptions and understanding about rubrics; seeking consensus in view of teacher's preferences in choosing the types of rubrics. In our work, we found that the following factors contributed to the challenges for our rubric engagement and collaborations:

Time: Look for, then protect, dedicated meeting and development time. When possible, we tried to meet for blocks of 1-2 hours.

Commitment: Even if potential collaborators have the time, they will also need to be committed to the project and committed to the other collaborators in their rubric team.

Team player: Each person must be willing (and able) to truly listen to the comments/insights of the others.

Similar values: As mentioned by Gabriel (2017), misalignment of values contributes to staff burnout. Having similar values as your fellow collaborators helps build trust among the team as you work together toward a common goal.

Peer led: By asking individual lecturers to first showcase their own rubrics or approaches to assessment, this created an engaging peer learning environment, where lecturers were hearing from peers who faced similar workload challenges and were able to provide practical examples of how the adoption of rubrics had benefitted them and their students.

At the early stage of the process, it was clear that some teams had developed more detailed rubrics than others. For example, the early stage presentation skills rubric (Figure 6) had clear grading categories but did not provide detail on what might be expected within each grading band. The team developing the Consultancy Project rubric (Figure 7), benefitted from previous work and there was possibly greater clarity of expectations. Earlier drafts included detailed descriptions within each grading band.

Figure 6. Early draft consultancy project rubric

Presentation Skills Rubric:

		Failed to meet expectations	Meets Expectations	Exceeds Expectations	Additional Comments
Content **Marks out of 30** **Weighting: 30%**	Originality (10)				
	Attention to brief (10)				
	Quality of supporting documentation (10)				
Oral Presentation (Effectiveness – informal, personal, immediate, active) **Marks out of 60** **Weighting: 60%**	Eye Contact (10)				
	Body Language (open) (10)				
	Voice (pitch, pace, pause, volume) (10)				

Figure 7. Early Draft Consultancy Project Rubric

Student Name		Form completed by:
Project Title		Please X as appropriate
Supervisor		
Second Reader		

To be completed by both Supervisor & Second readers and returned to the **Graduate Business School Office** by **(INSERT DATE)**
Ensure to select grade bracket for each section & provide comments/ feedback for each section.

	0-29% (Fail)	30-39% (Fail)	40-49% (Pass)	50-59% (2.2)	60-69% (2.1)	70-79% (First)	80%+ (First)
Explanation & justification of research topic & objectives 10%	Little or no academic rationale for topic or construction of research questions.	Failure to connect academic literature & research question &/or failure to pose proper research question	Some evidence of competence in selection & justification of topic area and research objectives.	Mostly competent selection & justification of topic and research objectives.	High competence shown in selection & justification but not always reaching high competence.	High competence in selection & justification of research objectives & topic but lacking flair.	Topic & research objectives of the quality & with the level of justification appearing in a quality academic peer-reviewed journal.
Insert X							
Comments:							

Furthermore, approaches to measuring PLOs vary amongst AACSB Schools. Sometimes measures of performance are taken within a course/module assessment and also contribute towards a course/module grade. In other cases the measure is taken at programme level (e.g. a test of general business knowledge which covers multiple modules). As such, instead of assigning a percentage grade, many AACSB Schools use three categories of Doesn't Meet Expectations/ Meets Expectations/ Exceed Expectations. As we were early in our discussions about our approach to measuring PLOs, the approach taken by the rubric working teams varied.

SOLUTIONS AND RECOMMENDATIONS

The following was agreed by the collaborative rubric teams on the Business college-level common rubrics, and were based on what they had learned through the literature they engaged with on rubric development and through the peer review process of the rubrics themselves.

Rubric Structure

- Identify around five key evaluative criteria.
- All the common rubrics will include descriptions of student performance at different levels of quality, which means that there is an inbuilt 'feed-forward' feature in all.
- Decide the scoring strategy for generating a final mark.

Rubric Language

- The need for a common language on 'Teamwork' was proposed and a decision was taken to proceed with this as a College-wide approach.
- Language in the rubric: we are using qualitative language in all the target statements.
- Avoid subjective language in favour of explicit guidance.

Rubric Use

- Rubrics can be used by programme teams formatively and summatively. Programme teams will be making decisions on 'within-module' vs 'programme level' measures of PLO attainment. While it is sometimes easier to use the 'Does not Meet/Meets/Exceeds Expectations' for certain PLOs - such as Teamwork, however, this means it would have to be scored and feedback given but it would not be part of a module grade. This entails extra workload for staff. On the flip side, some Schools have used a pre-test and post-test with questions from different course/module areas to test for knowledge of business etc, that is administered centrally (Tarnoff, 2023).
- The rubrics will encourage the students to use the feedback they receive. Co-constructing rubrics with students can effectively utilise the rubrics' potential to help students become reflective and thoughtful.
- We are avoiding designing overly complex rubrics and are focusing on the student's work providing evidence of meeting, exceeding, or falling short of the quality being looked for.
- Where multiple assessments within a programme have a similar structure or focus, using the same marking rubric is recommended.

Rubric Training

- Ensure that students have early opportunities to engage with the rubric.
- Continue to work as a team to master the rubric.

- Decide how the rubric will be used in the programme to provide feedback.

Rubric Supports

One key to the success of this initiative was how expanded, multi-disciplinary teams exploited the possibilities offered by digital technologies and collaborated to cultivate the practice of digitally-informed assessment design and development to create a high-quality learning experience for all students. We wanted to optimise our use of the VLE Brightspace - in using its rubric functionality, we would be making the move to paperless assignments and feedback. This would involve the use of purely digital rubrics embedded in the VLE and available to the students for their full learning journey. A number of colleagues had integrated some rubrics into Brightspace where they could take a paperless approach, develop the rubric quickly and also publish it so that students could access it when they logged in. Discussions were had about embedding all rubrics in Brightspace going forward so that students have a record of their learning, and performance against PLOs, so that they can be take greater ownership for their learning. As part of the suite of Rubric eResources, we proposed having a visual for each college-wide PLO alongside digital posters for promoting:

- Industry Engagement which highlights the applied nature of assignments, guest speakers, and work placement - citing numbers and linking to a new Faculty calendar.
- Preparing graduates for Career Success - highlighting graduate success stories. These should be discipline specific e.g. Accounting, Management, Marketing. They would feature high quality images of graduates with a quote and details on their new roles.
- Making a Difference - citing community work, work with charities and not-for-profits.

FUTURE RESEARCH DIRECTIONS

Following a period of organsiational design and restructuring, the Faculty have now set about conducting School and programmatic reviews, for every programme in the Faculty. This exercise presents an excellent opportunity to review the current PLOs and ensure that updated PLOs are adopted for every programme and more importantly that they are mapped across the programme and across modules so that programme teams, and students, have clarity with regard to where topics are introduced, reinforced and assessed. A final year assessment will be measured for AOL purposes with a feedback loop to ensure that programmes are continuously enhanced, in line with AACSB standards. Along with revising PLOs, part two of our rubrics project will be commenced, whereby existing rubrics will be reviewed, having been used with numerous student groups and by a broad range of lecturers. Additional PLOs will be developed in line with the new PLOs and all will be made available in a central repository to help to drive consistency in assessment and to further encourage students to take ownership of their learning journey.

During the second stage of the rubrics project, further consideration will be given to how rubrics might be utilised to assess related knowledge and skills but also to encourage authenticity, collaboration and peer assessment in an era of continuous and unabated technological advancement. Given the recent advances in Artificial Intelligence (AI), a number of workshops have been held with staff and UG and PG students to discuss how best to address concerns on ChatGPT and to leverage the benefits of these new technologies. Our study can be continued by exploring the pros and cons of leveraging such tools

to create rubrics as well as look at how rubrics might be used to mitigate against the use of AI in assessment - in particular, with the help of growth-oriented and skill-based rubrics, as argued by Lee (2023), peer assessment can encourage students to turn in original and self-crafted work rather than depending on AI to complete assignments.

CONCLUSION

As Technological University Dublin has recently entered into a new Faculty, School and Discipline structure, this initiative seeks to create a new way of collaborative working between cross-Faculty, multi-disciplinary teams. At its centre, it explores how to engage staff and students in the assessment change process utilising shared modules and rubrics. Underpinning our work is a social constructivist approach, where the use of rubrics are increasing learner activeness, curiosity, and engagement, all while reducing the workload of a lecturer, and co-designed and developed in a collaborative way. Through our upcoming programmatic reviews, the use of rubrics will allow students to engage better with deeper application of module concepts and meaningful demonstration of learning outcomes. Two major issues that programme teams have been facing since the return to in-person classes post-pandemic will guide the direction of our future work with collaborative college-wide rubrics. Firstly, student engagement - we are looking at ways to consider using rubrics as a vehicle for student engagement as a meaningful snapshot of student progress is more needed than ever. Secondly, over assessment - through the upcoming College-wide Programmatic Reviews, an opportunity exists to cut back on quantity of assessment across a programme, and use rubrics to steer this. We hope that this chapter will be a useful guide for those in a similar situations in higher education institutions in developing common rubrics with their cross-disciplinary programme teams.

REFERENCES

AACSB. (2019). Accreditation Standard 8 (2013 Business Standards): Curricula Management and Assurance of Learning An Interpretation, AACSB. https://www.aacsb.edu/insights/briefings/standard-8-white-paper

AACSB. (2020). 2020 Guiding Principles and Standards For Business Accreditation. https://www.aacsb.edu/educators/accreditation/business-accreditation/aacsb-business-accreditation-standards

Allen, S., & Knight, J. (2009). A Method for Collaboratively Developing and Validating a Rubric. *International Journal for the Scholarship of Teaching and Learning*, *3*(2), 3. doi:10.20429/ijsotl.2009.030210

Andrade, H., & Valtcheva, A. (2009). Promoting Learning and Achievement Through Self-Assessment. *Theory into Practice*, *48*(1), 12–19. doi:10.1080/00405840802577544

Andrade, H. G. (2005). Teaching with Rubrics: The Good, the Bad, and the Ugly. *College Teaching*, *53*(1), 27–30. doi:10.3200/CTCH.53.1.27-31

Brookhart, S. M. (2018). Appropriate criteria: Key to effective rubrics. *Frontiers in Education*, *3*, 22. doi:10.3389/feduc.2018.00022

Carless, D., & Boud, D. (2018). The development of student feedback literacy: Enabling uptake of feedback. *Assessment & Evaluation in Higher Education*, *43*(8), 1315–1325. doi:10.1080/02602938.2 018.1463354

Cuban, L. (1993). The lure of curriculum reform and its pitiful history. *Phi Delta Kappan*, *75*(2), 181–185.

Dawson, P. (2017). Assessment rubrics: Towards clearer and more replicable design, research and practice. *Assessment & Evaluation in Higher Education*, *42*(3), 347–360. doi:10.1080/02602938.2015.1111294

English, F. (1978). *Quality control in curriculum development*. American Association of School Administrators.

Francis, J. E. (2018). Linking Rubrics and Academic Performance: An Engagement Theory Perspective. *Journal of University Teaching & Learning Practice*, *15*(1), 5–22. https://ro.uow.edu.au/jutlp/vol15/ iss1/3. doi:10.53761/1.15.1.3

Gabriel, S. (October, 2017). Moving from Silos and Burnout to Community and Engagement. *Faculty Focus.* Higher Ed Teaching Strategies from Magna Publications.

Garrett, N., Marques, J., & Dhiman, S. (2012). Assessment of business programs: A review of two models. *Business Education & Accreditation*, *4*(2), 17–25.

Gibbs, G., & Dunbar-Goddet, H. (2009). Characterising programme-level assessment environments that support learning. *Assessment & Evaluation in Higher Education*, *34*(4), 481–489. doi:10.1080/02602930802071114

Gibbs, G., Rust, C., Jenkins, A., & Jacques, D. (1994). *Developing students' transferable skills*. The Oxford Centre for Staff Development.

Hale, J. A. (2008). *A guide to curriculum mapping*. Corwin Press.

Harland, T., & Wald, N. (2021). The assessment arms race and the evolution of a university's assessment practices. *Assessment & Evaluation in Higher Education*, *46*(1), 105–117. doi:10.1080/0260293 8.2020.1745753

Honeychurch, S. (2015). *Rubrics as a guide to student writing and staff grading*. Chartered Association of Business Schools.

Huball, H., & Burt, H. (2004). An Integrated Approach to Developing and Implementing Learning Centred Curricula. *The International Journal for Academic Development*, *9*(1), 51–65. doi:10.1080/1360144042000296053

Jacobs, H. (1997). *Mapping the big picture: Integrating curriculum and assessment K-12*. Association for Supervision and Curriculum Development.

Jessop, T., Hakim, Y., & Gibbs, G. (2014). The whole is greater than the sum of its parts: A large-scale study of students' learning in response to different programme assessment patterns. *Assessment & Evaluation in Higher Education*, *39*(1), 39. doi:10.1080/02602938.2013.792108

Jessop, T., & Tomas, C. (2016). The implications of programme assessment patterns for student learning. *Assessment & Evaluation in Higher Education*.

Jones, L., Allen, B., Dunn, P., & Brooker, L. (2017). Demystifying the rubric: A five-step pedagogy to improve student understanding and utilisation of marking criteria. *Higher Education Research & Development*, *36*(1), 129–142. doi:10.1080/07294360.2016.1177000

Jonsson, A., & Panadero, E. (2017). The Use and Design of Rubrics to Support AfL. In D. Carless, S. Bridges, C. Chan, & R. Glofcheski (Eds.), *Scaling Up Assessment for Learning in Higher Education* (pp. 99–111). Springer. doi:10.1007/978-981-10-3045-1_7

Joyner, H. S. (2016). Curriculum mapping: A method to assess and refine undergraduate degree programs. *Journal of Food Science Education*, *15*(3), 83–100. doi:10.1111/1541-4329.12086

Jutras, D. (2023). *Scales, stars and numbers: the question of evaluation*. Times Higher Education.

Lee, J. (2023). *Effective assessment practices for a ChatGPT-enabled world*. Times Higher Education.

Logan, D., King, J., & Fischer-Wright, H. (2011). *Tribal Leadership: Leveraging Natural Groups to Build a Thriving Organization*. Harper Business.

Morton, J. K., Northcote, M., Kilgour, P., & Jackson, W. A. (2021). Sharing the construction of assessment rubrics with students: A Model for collaborative rubric construction. *Journal of University Teaching & Learning Practice*, *18*(4). Advance online publication. doi:10.53761/1.18.4.9

Okojie, M. U., Bastas, M., & Fatma Miralay, F. (2022). Using Curriculum Mapping as a Tool to Match Student Learning Outcomes and Social Studies Curricula. *Frontiers in Psychology*, *13*, 850264. Advance online publication. doi:10.3389/fpsyg.2022.850264 PMID:36059751

Olson, J., & Krysiak, R. (2021). Rubrics as Tools for Effective Assessment of Student Learning and Program Quality. Curriculum Development and Online Instruction for the 21st Century (pp.173-200).

Panadero, E., & Jonsson, A. (2013). The use of scoring rubrics for formative assessment purposes revisited: A review. *Educational Research Review*, *9*, 129–144. doi:10.1016/j.edurev.2013.01.002

Panadero, E., & Jonsson, A. (2020). A critical review of the arguments against the use of rubrics. *Educational Research Review*, *30*, 30. doi:10.1016/j.edurev.2020.100329

Ragupathi, K., & Lee, A. (2020). Beyond Fairness and Consistency in Grading: The Role of Rubrics in Higher Education. In C. Sanger & N. Gleason (Eds.), *Diversity and Inclusion in Global Higher Education*. Palgrave Macmillan., doi:10.1007/978-981-15-1628-3_3

Rahmat, R. (2017). *An Approach to Overcome the Challenges of Using Rubrics to Improve Student's Understanding of Success Criteria*. World Association of Lesson Studies.

Reddy, Y. M., & Andrade, H. (2010). A review of rubric use in higher education. *Assessment & Evaluation in Higher Education*, *35*(4), 435–448. doi:10.1080/02602930902862859

Robley, W., Whittle, S., & Murdoch-Eaton, D. (2005a). Mapping generic skills curricula: A recommended methodology. *Journal of Further and Higher Education*, *29*(3), 221–231. doi:10.1080/03098770500166801

Robley, W., Whittle, S., & Murdoch-Eaton, D. (2005b). Mapping generic skills curricula: Outcomes and discussion. *Journal of Further and Higher Education*, *29*(4), 321–330. doi:10.1080/03098770500353342

Sadler, D. R. (2009). *Transforming holistic assessment and grading into a vehicle for complex learning.* Springer. doi:10.1007/978-1-4020-8905-3_4

Spencer, D., Riddle, M., & Knewstubb, B. (2012). Curriculum mapping to embed graduate capabilities. *Higher Education Research & Development*, *31*(2), 217–231. doi:10.1080/07294360.2011.554387

Stassen, M. L. A., Doherty, K., & Poe, M. (2011). *Course-based review and assessment: methods for understanding student learning.* Robert Langhorst & Company Booksellers.

Sumsion, J., & Goodfellow, J. (2004). Identifying generic skills through curriculum mapping: A critical evaluation. *Higher Education Research & Development*, *23*(3), 329–346. doi:10.1080/0729436042000235436

Tarnoff, K. (2023). How to Make Sure an AoL System Is Working. https://www.aacsb.edu/insights/articles/2023/06/how-to-make-sure-an-aol-system-is-working

APPENDIX: CONSULTANCY PROJECT RUBRIC

Table 4. Postgraduate Consultancy/ Applied Project Evaluation Form

Student Name		Form completed by: Please X as appropriate
Project Title		
Supervisor		
Second Reader		

To be completed by both Supervisor & Second readers and returned to the relevant School Office
Ensure to select grade bracket for each section & provide comments/ feedback for each section.

	0-29% (Fail)	30-39% (Fail)	40-49% (Pass)	50-59% (2.2)	60-69% (2.1)	70-79% (First)	80%+ (First)
Explanation & justification of Problem & Context **20%**	Poor explanation of problem. Little or no rationale for problem. Limited explanation of context.	Problem is explained but rationale is poor, failure to adequately connect problem to context.	Some evidence of competence in selection & justification of problem to be addressed within appropriate context.	Mostly competent selection & justification of problem to be addressed within appropriate context.	High competence shown in selection & justification of problem to be addressed within appropriate context but not always high competence.	High competence shown in selection & justification of problem to be addressed within appropriate context but lacking flair.	Problem is explained and justified to a very high standard (to the standard appearing in quality industry report/ article), problem is embedded within context.
Insert X							
Comments:							
Literature Review **20%**	Little evidence of reading or understanding alternative viewpoints.	Limited awareness of literature. Little independent understanding.	Evidence of understanding of literature, little critical analysis.	Mostly critical & adequate coverage of literature but significant lack of coverage or interpretation in some areas.	Overall a comprehensive critical analysis of literature but some over-simplification or limited interpretation.	Excellent use of appropriate literature in terms of identifying, evaluating & critically analyzing relevant theories, ideas & frameworks	Excellent & creative use of appropriate literature identifying, evaluating & critically analyzing relevant theories, ideas & frameworks.
Insert X							
Comments:							

	0-29% (Fail)	30-39% (Fail)	40-49% (Pass)	50-59% (2.2)	60-69% (2.1)	70-79% (First)	80%+ (First)
Research Methodology **15%**	Poor explanation and lack of justification of method.	Weak explanation and justification of method.	Some evidence of explanation and justification of method but inadequate.	Mostly adequate explanation and justification of method & process, but some significant flaws.	Adequate explanation and justification of method & process, with few lapses.	Concise explanation and justification of method & process, linked to problem.	Insightful and creative explanation and justification of method & process, strongly linked to problem.
Insert X							
Comments:							
Analysis of Findings **20%**	Little or no analysis of primary research.	Fails to demonstrate adequate competence in analysis.	Shows some competence of analysis in some areas	Shows clear competence of analysis but fails to adequately link to context and theory	Competent analysis of primary research within context and theory but lacks greater reflection & sophistication	Critical & insightful interpretation of primary research within context and theory	Insightful & creative interpretation & presentation. Rigorous links of evidence, within context and theory
Insert X							
Comments:							
Recommendations **25%**	Few or no recommendations made.	Shows need to make findings relevant to appropriate audiences. Weak recommendations	Logical recommendations made, failure to adequately reflect on context & theoretical framework.	Concise recommendations made, somewhat reflective of context and theoretical framework.	Concise recommendations made, largely reflective of context and theoretical framework.	Largely compelling recommendations made. Demonstrates critical thinking and reflection.	Clear, detailed, evidence based recommendations made. Demonstrates critical thinking and reflection.
Insert X							

Continued on following page

Table Continued

Comments:
Additional Comments (if applicable):

Awarded Mark: Third Reader (if required):

Supervisor	
Second Reader	
Agreed	
Date	
Mark	
Date	

In addition to the sections detailed above, also consider the following when determining an overall mark for the project:

Structure and flow of project, connectivity of chapters.

Presentation of report.

Referencing.

Chapter 10
Developing a Rubric to Evaluate the Dissertations Conducted in the Fields of Educational and Social Sciences

Ömer Açıkgöz
Social Sciences University of Ankara, Turkey

Aydın Aslan
Selcuk University, Turkey

Korkut Koçak
Republic of Türkiye Ministry of National Education, Turkey

Aslı Günay
Social Sciences University of Ankara, Turkey

Nevzat Yavuz
The Scientific and Technological Research Council of Türkiye, Turkey

ABSTRACT

This study aims to develop a rubric to evaluate the dissertations implemented in the fields of educational and social sciences. In the development of this rubric, the acquisition requirements concerning knowledge, skills, and competence at the doctorate level in the European and Turkish qualifications frameworks, the legal framework of Turkish higher education, and the perceptions of 12 experts in the fields of educational and social sciences concerning the common competences of the dissertations were considered. The rubric can contribute to the evaluation of dissertations completed in the field of educational and social sciences concerning these dimensions and provide PhD students, researchers, and academics with a guide to evaluate their academic studies based on an empirical instrument.

INTRODUCTION

A Doctor of Philosophy (Ph.D.) dissertation is the final and highest educational outcome in higher education. Hall (1991) indicated that the academic significance of a discipline depends on the quality of doctorate studies, in which Ph.D. dissertations are seen as important tools for a discipline to improve academically. Briefly, Ph.D. dissertation demonstrates its author's technical, analytical, and writing skills (Lovitts, 2005). The doctorate level is the highest qualification level in the European Qualifications Framework (EQF) and the Turkish Qualifications Framework (TQF), which refers to level 8 and usually requires four years of study, mostly as a period of research. Ph.D. holders are expected to confirm their advanced knowledge, skills, and competencies in their dissertations according to their qualification level description (EQF, 2015; TQF, 2016). Besides, national authorities can take additional measurements to improve the quality of the doctorate programs and Ph.D. dissertations. For example, the Turkish Graduate Education Regulation (TGER) in 2016 emphasizes that the Ph.D. holders in Türkiye are required to meet one of the conditions, namely bringing innovation to science, developing a new scientific method or implementing a known method to another field (TGER, 2016). In this regard, they are expected to contribute to the literature or the application through their dissertations. They are important scientific studies with findings, implications, and suggestions attracting various researchers, policymakers, and academics.

Several questions appear in the evaluation of a dissertation: what determines a basic dissertation with minimum requirements? Who decides what characteristics it should have? These are the critical questions, which should be considered in the evaluation of Ph.D. dissertations. In particular, although countries have different higher education systems and regulations, most universities have issued some general or specific guidelines on quality and standards for Ph.D. dissertations evaluations regardless of the locations. For instance, properties like originality, sound methods, significant contribution to knowledge, and publishable results are common worldwide criteria (Kyvik & Thune, 2015).

While the TGER establishes a standard framework for evaluating Ph.D. dissertations in Türkiye, each university in Türkiye has its own regulations for Ph.D. dissertations processes, regardless of the academic field. The dissertation jury consists of five members: three faculty members, including the student's dissertation supervisor, and at least two representatives from other higher education institutions. Moreover, evaluation processes might differ according to the students' and supervisors' characteristics, experiences, and relations. Hence, it is not possible to say that they have particular criteria to evaluate dissertations in Türkiye. For that reason, this study tries to fill this gap in the literature by developing a rubric through which jury members, academics, researchers and doctoral students can evaluate the dissertations conducted in the fields of educational and social sciences. Furthermore, the dimensions in the rubric can provide them with a profound and comprehensive idea about the quality of the Ph.D. dissertations.

Developing a rubric to evaluate Ph.D. dissertations conducted in the fields of educational and social sciences based on the doctorate qualifications, based on knowledge, skills, and competencies descriptors can contribute to the literature. Regarding this, qualification descriptors of the EQF and TQF were used, providing a comprehensive overview of the quality of doctorate programs and Ph.D. dissertations in European countries and Türkiye for the possible acquiring learning outcomes during their studies. In this regard, universities in both Türkiye and European countries might use the newly developed rubric by this study as a reference source to evaluate the quality of Ph.D. dissertations.

BACKGROUND

The rubric, which is often used in dissertation evaluations, is one method of evaluating dissertations. It is a scoring tool used in qualitative rating to evaluate original or complex student works. It consists of criteria for rating significant dimensions of performance and standards to what extent those criteria are attained (Jonsson & Svingby, 2007). It promotes peer assessment, self-assessment, and academics assessment (Andrade, 2005). In this regard, academics, educators, and researchers can use it with higher education students. There are three fundamental features in a rubric: evaluation criteria, quality dimensions, and a scoring strategy (Popham, 1997). It can decrease rater biases with its construct consisting of criteria and performance levels and provide students with more realistic and detailed feedback concerning their performance (Parlak & Doğan, 2014).

The rubric can be holistic or analytic, depending on its purpose of usage. Scoring in a holistic rubric is implemented based on students' overall impressions. It gives evaluators a quick scoring and overall impression of students' performance on a task despite not involving detailed feedback. On the other hand, scoring in an analytic rubric is conducted according to several dimensions. Therefore, evaluators can get more detailed feedback and consistent scoring across students and graders though it is time-consuming to score (Zimmaro, 2004). Another important distinction about the rubric is that it can be general and task-specific. A general rubric is used when all performances and tasks are evaluated with the same rubric. If a particular performance or task is aimed to be evaluated, a task-specific rubric is employed (Özen, 2019). The purpose of usage determines whether a rubric will be holistic, analytic, general or task-specific.

Rubrics are employed in many disciplines such as the liberal arts, management, and teacher education (Reddy & Andrade, 2010). For example, a rubric was developed to grade undergraduate students' APA-style introductions (Stellmack, Konheim-Kalkstein, Manor, Massey & Schmitz, 2009). Rubrics are consulted to evaluate doctoral dissertations (Agu, Omenyi & Odimegwu, 2015; Fitt, Wlaker, Leary, 2009; Lovitts, 2005). One reason to use rubrics in the evaluation of doctoral dissertations is that they make the evaluation criteria more explicit (Johnsson & Svingby, 2007). Therefore, several universities (Texas A & M Commerce University, 2016) utilize them to evaluate theses and dissertations.

In this context, the main objective of this study is to develop a rubric to evaluate the Ph.D. dissertations implemented in the fields of educational and social sciences. Hence, standard criteria might be defined for the evaluation processes of Ph.D. dissertations to improve their quality. As the qualifications in the EQF, TQF, and TGER consist of several dimensions, an analytic rubric was embraced to elicit more detailed feedback for the dissertations completed in the field of educational and social sciences with regard to theoretical framework, method, and contributions of research dimensions. The knowledge, skills, and competencies of researchers can be measured in their dissertations with this rubric. For this purpose, 20 dissertations were randomly selected from the National Theses Center of the Council of Higher Education database in Türkiye, 10 in the field of educational sciences and 10 in the field of social sciences. Besides, 12 experts in different educational and social sciences departments in Turkish universities (two experts in measurement and evaluation in educational sciences, one professor in educational administration, two professors in educational sciences, one associate professor in psychology, one associate professor in sociology, one associate professor in mathematics, one assistant professor in economy, one assistant professor in international relations, one professor in economy, and one expert in public administration) were asked to define the common criteria for Ph.D. dissertations to lessen the quality differences among them in the fields of educational and social sciences.

METHODOLOGY

In developing the current rubric, the dimensions with regard to the EQF, TQF, and TGER were considered. Besides, the studies conducted by Karslı, Karabey, Çağıltay, and Göktaş (2018) and Lovitts (2005) were taken into account qualification dimensions to determine the items in the rubric. The perceptions of 12 experts in their fields concerning their academic standards were also taken into account to form the items in the rubric. The rubric consists of three dimensions "Theoretical Framework", "Method" and "Contributions of Research (AppendixA). The first one is related to the clarity of the scientific importance, research problem, and the theory/theories on which the dissertation is based. The second one is closely related to scientific knowledge and appropriate methodologies, including the notions of validity, reliability, correctness, truthfulness, and consistency. The last one indicates originality, scientific contributions to knowledge, publishable results, and practical and socio-economic assistance.

After all, these dimensions were formed based on the EQF and TQF, TGER, the literature review and the perceptions of 12 experts in the field of educational, and social sciences. Hence, there are 11 items in the rubric, which describe these dimensions. The items are rated in accordance with "no", "partial" and "yes". A three-category grading was chosen in this rubric because of two reasons. The first reason is that a gradual and hierarchic three-category grading in the rubric makes the criteria of the rubric more understandable and explicit. The second one is that it was decided that using a three-category grading would be more appropriate based on the perceptions of 22 experts (seven experts in the field of measurement and evaluation, one associate professor, and three assistant professors in the field of educational administration, two assistant professors in the field of educational sciences, two associate professors and one assistant professor in curriculum and instruction, one associate professor in sociology, one assistant professor in psychology, one professor and one assistant professor in the economy, one expert in the field of public administration and one assistant professor in the international relations) concerning the items in the rubric.

In the development of the rubric, the perceptions of 22 experts considered. It was paid attention to select the experts with the theoretical and practical knowledge to be able to supply more comprehensive and profound feedback for the rubric. 15 in 22 experts have been experienced to be a jury member in master's thesis and doctoral dissertation defenses.

The researchers of this study in collaboration with an independent researcher carefully examined the dimensions in the EQF, TQF, and TGER, in addition to the literature review and 12 experts' perceptions concerning the common criteria for the high-quality dissertations to determine which items should be included under each dimension in the rubric. In this context, 15 tentative items were prepared and sent to 22 experts. These experts were sent an evaluation form to indicate their perceptions of the items and were asked to score the tentative 15 items in the rubric as "Acceptable", "Revised" and "Not acceptable".

In the context of the content validity of the tentative data collection instrument, Content Validity Ratio (CVR) and Content Validity Index (CVI) based on expert views (Lawshe 1975, 568) were used to determine which items should be omitted in the rubric. In the determination of experts, who have adequate knowledge and enough time to evaluate a data collection instrument, at least 5 experts and at most 40 experts can take part in an evaluation of the instrument (Yurdugül 2005, 2). In the current rubric development, the perceptions of 22 experts were derived in the calculation of CVR for each item and CVI for the rubric and its dimensions. CVR is calculated in the following way (Lawshe 1975, 567):

$$CVR = N_E - (N/2) / N/2$$

N_E shows the number of experts indicating an item "Acceptable" and N the number of experts in this formula. The items with 0 (zero) or lower than this value are directly omitted from the tentative instrument and Content Validity Criterion (CVC) is calculated for each remaining item at the level of significance (α=.05). CVC (CVRcritical) is a minimum CVR value required to remove the possibility to find each item appropriate in the instrument by chance and make correct decisions for items.

FINDINGS

According to CVRcritical the table prepared by Ayre and Scally (2014) in consideration of Lawshe's study, the minimum CVR value should be at least 0.455 for one item 22 experts examine at the level α=0.05 (type 1 error). According to Table 1, 11 in 15 items met the accepted conditions (CVR ≥ CVR critical (N=22. α=0.05)=.455).

Table 1. CVR values for the tentative items

Item	CVR		Item	CVR
Item1	1,000		Item9	0.158*
Item2	0,895		Item10	0.579
Item3	0,789		Item11	0.684
Item4	0,684		Item12	-0.158*
Item5	0,895		Item13	0,789
Item6	0,789		Item14	0.684
Item7	0,053*		Item15	0.368*
Item8	0,895			
*CVR<CVR$_{Critical}$(N=22, α=.05)				

Content Validity Index (CVI) is calculated for all tests and if available its sub-dimensions after CVR treatment to include items in the final test. CVI critical value is the same as CVR, which is valid for items (Yurdugül, 2005). So, the minimum CVI value for the rubric and its sub-dimensions is .455 for 22 experts at the level of α=0.05 (type 1 error).

Table 2. CVI values for all rubric and its sub-dimensions

Rubric and its sub-dimensions		CVI
The Whole/entire rubric		.789*
Theoretical framework		.895*
Method		.789*
Contribution of research		.726*
*CVI≥CVI$_{critical}$(N=22, α=.05)		

According to Table 2, CVI values for the rubric and its sub-dimensions are higher than CVI critical value (CVI ≥ CVcritical (N=22, α=0.05)=.455). Based on CVR and CVI values, the content validity of the rubric is acceptable.

Besides, the construct validity of the rubric was ensured through the perceptions of these experts to evaluate how well the rubric measures dissertations with regard to theoretical framework, method and contribution of research dimensions. Its reliability was checked through the rater reliability because this type of reliability is the most frequently considered type of reliability in rubric development (Moskal & Leydens, 2000). So, 20 dissertations were randomly selected from the National Theses Center of the Council of Higher Education database, 10 in the field of educational sciences and 10 in the field of social sciences. Six experts in their disciplines, three in educational sciences and three in social sciences assessed them. All the experts were informed that the researchers aimed to use a three-point scale namely "no", "partial" and "yes", which corresponds to 0, 1, and 2 in the rubric, respectively. The directions to score the items in the rubric were provided with the rubric (See AppendixB). To ensure the reliability of the rating, it was checked whether each expert understood the rating procedure properly. They were asked to evaluate the dissertation through the rubric by choosing "no", "partial" and "yes" on a scale in the rubric. Krippendorff's Alpha and Fleiss' Kappa values were calculated to examine the consistency of assessments among the raters in this study. Krippendorff's α was calculated through a macro produced for the SPSS package program by Hayes and Krippendorff (2007).

Table 3. Krippendorff's alpha values for the rubric development in the field of educational sciences

A	95% CI		Units	Raters	Pairs
	Lower	Upper			
.912	.851	.9712	10	3	30

As seen in Table 3, Krippendorff's alpha values for the dissertations implemented in the field of educational sciences were found to be .912. This value is acceptable [α>0.667] (Krippendorf, 2004). Fleiss' Kappa value for these dissertations was calculated to support the former result. According to Table 4, Fleiss' kappa values [κ=.703, p < .05] show a substantial agreement among the raters (Landis & Koch, 1977). The rubric can be regarded as sufficiently acceptable as a result of these calculations.

Table 4. Fleiss kappa values for the rubric development in the field of educational sciences

K	Asymptotic S.E.	z	p	95% Asymptotic CI Bound	
				Lower	Upper
.703	.118	5,951	.000	.471	.935

On the other hand, Krippendorff's alpha values and Fleiss Kappa values were also calculated for the dissertations implemented in the field of social sciences.

Table 5. Krippendorff's alpha values for the rubric development in the field of social sciences

A	95% CI		Units	Raters	Pairs
	Lower	Upper			
.952	.932	.970	10	3	30

As seen in Table 5, Krippendorff's alpha values were found to be .952. This value is acceptable [α>0.667] (Krippendorf, 2004). Fleiss' Kappa value was calculated to support the former result.

Table 6. Fleiss kappa values for the rubric development in the field of social sciences

K	Asymptotic S.E.	z	p	95% Asymptotic CI Bound	
				Lower	Upper
.341	.071	4.819	.000	.202	.480

According to Table 6, Fleiss' kappa values [κ=.703, p < .05] show a substantial agreement among the raters (Landis & Koch, 1977). The rubric can be regarded as sufficiently acceptable as a result of these calculations. So, it can be used in the evaluation of Ph.D. dissertations completed in the fields of educational and social sciences.

FUTURE RESEARCH DIRECTIONS

There are several limitations that should be considered dealing with the findings of this study. The current rubric was developed to evaluate Ph.D. dissertations in educational and social sciences. To use the rubric in other disciplines such as medical or engineering sciences, it may be necessary to ensure its validity and reliability. Also, this new rubric was developed mostly based on TQF and TGER descriptors; hence, not taking into account other countries' qualifications frameworks may be a shortcoming of this study. Since comparative studies were not done by experts working in different countries, it is difficult for us to know whether using the same criteria in examining dissertations is appropriate.

CONCLUSION

This study aims to determine the standard criteria for the evaluation processes of Ph.D. dissertations to improve their quality and lessen the differences between the current assessment procedures. Therefore, the new rubric was developed to evaluate the PhD dissertations conducted in the fields of educational and social sciences with regard to the EQF, TQF, TGER, the literature review, and the perceptions of 12 experts in their fields in Türkiye.

To ensure the rubric's reliability, the interrater agreement among the raters in the fields of educational and social sciences was checked separately. To identify the agreement, both Krippendorff's α and Fleiss'

kappa values among the three raters with a Ph.D. degree in educational sciences were calculated. Krippendorff's Alpha was 0.912 and Fleiss' Kappa was 0.703 for the educational dissertations. On the other hand, three raters in the field of social sciences rated 10 social sciences dissertations independently. Krippendorff's Alpha was 0.952 and Fleiss' Kappa was 0.341 for these dissertations. According to the review of 75 studies on rubric conducted by Jonsson and Svingby (2007), it was revealed that many estimates did not reach the criterion of 70% or greater. Stemler (2004) argued that 70% or greater is required for the exact agreement. Two types of interrater reliability methods were employed to enable the reliability of the current rubric. In this regard, it can be said that it met the conditions, which are necessary steps in a rubric development process with regard to reliability. As a result, it is seen that the interrater reliability of the current rubric is acceptable enough.

As in the case of the reliability of the rubric, not only content but also construct validity was implemented by consulting the experts' perceptions. The content validity ratio and content validity index values were calculated to ensure the content validity. These results indicated that it is acceptable with regard to content validity.

The rubric consists of 11 items with three dimensions: theoretical framework, method, and research contributions. There are three items in the theoretical framework, three items in the method, and five items in the contributions of research that describe different criteria for the high quality of dissertations. During the evaluation process of dissertations, these dimensions might be used as a whole or separately in practice.

The rubric appeals to all Ph.D. students, academics, supervisors, faculty, and external jury members. While doctorate students prepare their proposals for their dissertations, they can use it as a guideline. From this perspective, the rubric could have an impact on researchers to produce high-quality dissertations, which have high-value-added, sound methodology and wide impact criteria. In addition, supervisors can guide their students to prepare their dissertations regarding the items in the rubric. Faculty and external jury members can evaluate dissertations based on the common criteria in the rubric. Furthermore, the rubric can indicate to what extent Ph.D. holders are qualified according to each dimension of the rubric. This can provide policymakers and researchers with empirical and sound findings to improve doctoral training in their countries.

The newly developed rubric in this study is expected to enable the doctorate students in the field of educational and social sciences to prepare their Ph.D. dissertations at least the average standards. Hereby, doctorate students can conduct their dissertation studies with regard to the items related to the three determined dimensions: theoretical framework, method, and their studies' contributions to the field. So, the newly developed rubric can promote doctorate students to assess themselves. In addition, it can enable jury members to assess dissertations in the light of the standard criteria rather than their perceived assumptions. Also, this rubric might give us some information about to what extent doctorate students acquire the qualifications defined in TQF and EQF. In sum, the rubric might be used to improve the quality of dissertations. Thus, it might be possible to measure or monitor the quality of dissertations in a meaningful way at an aggregate level. To this end, this study might provide decision-makers with reference to improve the quality of doctorate programs around the world.

DISCUSSION

The rubric can be used as a guideline to enhance educational and social sciences doctoral training programs with respect to curriculum and graduates' competencies based on the concrete outputs of programs, namely Ph.D. dissertations. Academics and policymakers can determine the strengths and weaknesses of the programs with regard to the theoretical framework, method and contributions of research dimensions. To illustrate, when academics, namely jury members observe problems stemming from students' theoretical knowledge in their dissertations, they can make improvements in the curriculum, thereby making them take more theoretical courses in the program. With the improvements made in the programs, it can be possible to train more competent researchers in their study fields.

REFERENCES

Agu, N. N., Omenyi, A. S., & O, C. (2015). Evaluation of doctorate dissertation in Nigerian Universities: do faculties provide and use criteria / rubrics? *International Journal of Technology and Inclusive Education (IJTIE), 4*(1), 565-569.

Andrade, H. G. (2005). Teaching with rubrics: The good, the bad, and the ugly. *College Teaching*, *53*(1), 27–30. doi:10.3200/CTCH.53.1.27-31

Ayre, C., & Scally, A. J. (2014). Critical values for Lawshe's content validity ratio: Revisiting the original methods of calculation. *Measurement & Evaluation in Counseling & Development*, *47*(1), 79–86. doi:10.1177/0748175613513808

EQF. (2005). *Descriptors defining levels in the European Qualifications Framework*. EQF. https://ec.europa.eu/ploteus/tr/node/1440

Fitt, M. H., Walker, A. E., & Leary, H. M. (2009). Assessing the quality of doctoral dissertation literature reviews in instructional technology. *Paper presented at the Annual meeting of the American Educational Research Association San Diego*, 3-15.

Hall, C. M. (1991). Tourism as the Subject of Post-graduate Dissertations in Australia. *Annals of Tourism Research*, *18*(3), 520–523. doi:10.1016/0160-7383(91)90061-F

Hayes, A. F., & Krippendorff, K. (2007). Answering the call for a standard reliability measure for coding data. *Communication Methods and Measures*, *1*(1), 77–89. doi:10.1080/19312450709336664

Jonsson, A., & Svingby, G. (2007). The use of scoring rubrics: Reliability, validity and educational consequences. *Educational Research Review*, *2*(2), 130–144. doi:10.1016/j.edurev.2007.05.002

Karslı, M. B., Karabey, S. Ç., Nergiz, E., & Goktas, Y. (2018). Comparison of the discussion sections of PhD dissertations in educational technology: The case of Turkey and the USA. *Scientometrics*.

Kirk, R. E. (2007). *Statistics, an introduction* (5th ed.). Thomson Wadsworth.

Krippendorff, K. (2004). Reliability in content analysis: Some common misconceptions and recommendations. *Human Communication Research*, *30*(3), 411–433. doi:10.1111/j.1468-2958.2004.tb00738.x

Kyvik, S., & Thune, T. (2015). Assessing the quality of PhD dissertations. A survey of external committee members. *Assessment & Evaluation in Higher Education*, *40*(5), 768–782. doi:10.1080/02602938.2014.956283

Landis, J. R., & Koch, G. G. (1977). The measurement of observer agreement for categorical data. *Biometrics*, *33*(1), 159–174. doi:10.2307/2529310 PMID:843571

Lawshe, C. H. (1975). A quantitative approach to content validity. *Personnel Psychology*, *28*(4), 563–575. doi:10.1111/j.1744-6570.1975.tb01393.x

Lovitts, B. E. (2005). How to grade a dissertation. *Academe, 91*(6), 18-23.

Moskal, B. M., & Leydens, J. A. (2000). Scoring rubric development: Validity and reliability. *Practical Assessment, Research & Evaluation*, *7*(10), 1–6.

MYK. (2015). *Turkish Qualifications Framework*. MYK. https://www.myk.gov.tr//TRR/File6.pdf

Özen, S. O. (2019). *Öğrenenlerin e-değerlendirmeye dayalı kişiselleştirilmiş geri bildirim yollarının aranması [An investigation of the learners' personalized feedback paths based on e-assessment.]* [Unpublished PhD dissertation, Eskişehir Osmangazi Univeristy, Eskişehir, Turkey].

Parlak, B., & Doğan, N. (2014). Dereceli puanlama anahtarı ve puanlama anahtarından elde edilen puanların uyum düzeyleri [Comparison of answer key and scoring rubric for the evaluation of the student performance]. *Hacettepe Üniversitesi Eğitim Fakültesi Dergisi / H.U. Journal of Education*, *29*(2), 189–197.

Popham, W. J. (1997). What's wrong-and what's right-with rubrics? *Educational Leadership*, *55*(2), 72–75.

Reddy, Y. M., & Andrade, H. (2010). A review of rubric use in higher education. *Assessment & Evaluation in Higher Education*, *35*(4), 435–448. doi:10.1080/02602930902862859

Stellmack, M. A., Konheim-Kalkstein, Y. L., Manor, J. E., Massey, A. R., & Schmitz, J. A. P. (2009). An assessment of reliability and validity of a rubric for grading APA-style introductions. *Teaching of Psychology*, *36*(2), 102–107. doi:10.1080/00986280902739776

Stemler, S. E. (2004). A comparison of consensus, consistency and measurement approaches to estimating interrater reliability. *Practical Assessment, Research & Evaluation*, 9.

Texas A&M Commerce University (2016). *Rubrics for assessing dissertations*. Texas A & M Commerce University. https://www.tamuc.edu/academics/graduateschool/Thesis%20and%20Dissertation%20Services/graduate%20rep%20rubrics.aspx

TGER (2016). Turkish Graduate Education Regulation. *27561 numbered Official Gazette*.

Yurdugül, H. (2005). *Ölçek Geliştirme Çalışmalarında Kapsam Geçerliği için Kapsam Geçerlik İndekslerinin Kullanılması*. Paper presented at XIV Ulusal Eğitim Bilimleri Kongresi, Pamukkale Üniversitesi Eğitim Fakültesi, Denizli. http://yunus.hacettepe.edu.tr/~yurdugul/3/indir/PamukkaleBildiri.pdf

Zimmaro, D. M. (2004). *Developing grading rubrics*. BSU. http://bsuenglish101.pbworks.com/f/rubricshandout.pdf

ADDITIONAL READINGS

Arı, G. (2021). An overview of writing rubrics in doctoral dissertations in Turkey. *International Journal of Progressive Education*, *17*(2), 385–405. doi:10.29329/ijpe.2021.332.24

Maki, P. L., & Borkowski, N. A. (Eds.). (2006). *The assessment of doctoral education. Emerging criteria and new models for improving outcomes*. Stylus Publishing.

Panadero, E., & Jonsson, A. (2013). The use of scoring rubrics for formative assessment purposes revisited: A review. *Educational Research Review*, *9*(0), 129–144. doi:10.1016/j.edurev.2013.01.002

Rakedzon, T., & Baram-Tsabari, A. (2017). To make a long story short: A rubric for assessing graduate students' academic and popular science writing skills. *Assessing Writing*, *32*, 28–42. doi:10.1016/j.asw.2016.12.004

Rivas, M. A., De La Serna, M. C., & Martínez-Figueira, E. (2014). Electronic rubrics to assess competences in ICT subjects. *European Educational Research Journal*, *13*(5), 584–594. doi:10.2304/eerj.2014.13.5.584

The University of North Carolina at Chapel Hill. (2017). *Using rubrics to assess student learning outcomes at the program level.* UNC. https://oira.unc.edu/wp-content/uploads/sites/297/2017/07/Developing-and-Using-Rubrics.pdf

Vaccari, D., & Thangam, S. (2010). *A proposed doctoral assessment procedure and rubric for science and engineering. American Society for Engineering Education* Paper presented at 2010 Annual Conference & Exposition, Louisville, Kentucky. 10.18260/1-2--16106

Yağız, O., Ötügen, R., Kaya, F., & Aydın, B. (2014). A literature review analysis of the Turkish scholars' research articles in ELT and applied linguistics. *Procedia: Social and Behavioral Sciences*, *158*, 389–393. doi:10.1016/j.sbspro.2014.12.105

KEY TERMS AND DEFINITIONS

Dissertation: An essay or thesis written on a particular subject by a candidate to earn the doctorate degree.

Doctor of Philosophy: It refers to the highest degree in an academic level for graduate study.

Educational Sciences: It is any branch of academic study to describe, understand and prescribe the policies and practices to be developed in education.

European Qualification Framework: It is a common European reference framework to make qualifications more readable and understandable across different countries and systems.

Qualifications: They refer to the certificates and diplomas awarded following teaching and learning processes.

Rubric: It is a scoring tool that is used in qualitative rating to evaluate original or complex student works.

Social Sciences: It is any branch of academic study or science that deals with human behaviour in its social and cultural aspects.

Turkish Qualification Framework: It refers to the national qualifications framework which was prepared in line withthe European Qualifications Framework (EQF).

APPENDIX A

Table 7. Graded category rating scale

Dimensions	Item	No	Partial	Yes
THEORATICAL FRAMEWORK	1. The scientific importance of the dissertation was expressed clearly and straightforwardly.			
	2. The research problem of the dissertation was expressed clearly and straightforwardly.			
	3. The theory / theories, on which the dissertation is based, were expressed clearly and straightforwardly.			
METHOD	4. The study group / sample of the dissertation were appropriately determined in line with its aims.			
	5. The validity and reliability treatments of the dissertation were conducted in line with its aims.			
	6. The appropriate methods and techniques in the dissertation were used in line with its aims.			
CONTRIBUTIONS OF RESEARCH	7. The dissertation provided new and authentic contributions to the current literature or applications.			
	8. The dissertation provided new and authentic suggestions to the current literature or applications.			
	9. The dissertation has publication potential.			
	10. The dissertation makes contributions to one of economic, social and cultural fields.			
	11. The dissertation results make associations among different disciplines.			

APPENDIX B

Table 8. Rubric to develop dissertations conducted in educational and social sciences

Dimensions	Item	No	Partial	Yes
THEORATICAL FRAMEWORK	1. The scientific importance of the dissertation was expressed clearly and straightforwardly.	* The scientific importance of the dissertation was not expressed. * Supportive evidences with regard to the scientific importance of the dissertation were not provided.	* The scientific importance of the dissertation was expressed. However, it was not expressed clearly and straightforwardly * Weak evidences with regard to the scientific importance of the dissertation were provided.	* The scientific importance of the dissertation was expressed clearly and straightforwardly. * Strong and persuasive evidences with regard to the scientific importance of the dissertation were provided.
	2. The research problem of the dissertation was expressed clearly and straightforwardly.	* The research problem of the dissertation was not expressed. * The research problem was not supported with evidences. * The groups or cases affected from the research problem were not expressed.	* The research problem of the dissertation was expressed. However, it was not expressed clearly and straightforwardly. * The research problem was supported with weak evidence. * The groups or cases affected from the research problem were ambiguously expressed.	* The research problem of the dissertation was expressed clearly and straightforwardly. * The research problem was supported with strong and empirical evidences. * The groups or cases affected from the research problem were clearly expressed.
	3. The theory / theories, on which the dissertation is based, were expressed clearly and straightforwardly.	* The theory / theories, on which the dissertation is based, were not expressed. * The theory / theories which are not directly related with the dissertation were mostly expressed. * The contradictory or wrong information concerning the theory / theories, on which the dissertation is based, were provided.	* The theory / theories, on which the dissertation is based, were provided. However, it was not expressed clearly and straightforwardly. * The theory / theories, on which the dissertation is based, were not presented in a logical order and holistically. * The contradictory or wrong information concerning the theory / theories, on which the dissertation is based, were provided by mistake.	* The theory / theories, on which the dissertation is based, were expressed clearly and straightforwardly. * The theory / theories, on which the dissertation is based, were presented in a logical order and holistically. * The contradictory or wrong information concerning the theory / theories, on which the dissertation is based, were not provided.
METHOD	4. The study group / sample of the dissertation were appropriately determined in line with its aims.	* The method to determine the study group / sample of the dissertation was not expressed. * The method to determine the study group / sample of the dissertation was not consistent with the study subject and method of the dissertation. * Even though the sample of the study was determined, the method (s) used to determine the sample size was not expressed.	<u>Only</u> one point is met: * The method to determine the study group / sample of the dissertation was expressed. * The method to determine the study group / sample of the dissertation was consistent with the study.	* The method to determine the study group / sample of the dissertation was expressed. * The method to determine the study group / sample of the dissertation was consistent with the study subject and method of the dissertation. * The method (s) used to determine the sample size and sample of the study was expressed.
	5. The validity and reliability treatments of the dissertation were conducted in line with its aims.	* Appropriate evidences for the validity of the dissertation were not provided. *Appropriate evidences for the reliability of the data used in the dissertation were not provided. * If a measurement instrument was used in the dissertation; - The validity and reliability treatments of the measurement instrument were not conducted. - If the measurement instrument was developed beforehand, the information concerning its validity and reliability was not reported.	* Weak evidences for the validity of the dissertation were provided. * Weak evidences for the reliability of the data used in the dissertation were provided. * If a measurement instrument was used in the dissertation, its validity and reliability treatments were conducted. However, they did not meet the expectations.	* Appropriate and strong evidences for the validity of the dissertation were provided. * Appropriate and strong evidences for the reliability of the data used in the dissertation were provided. * If a measurement instrument was used in the dissertation; - Its validity and reliability treatments were appropriately conducted. - If the measurement instrument was developed beforehand, the information concerning its validity and reliability was reported.
	6. The appropriate methods and techniques in the dissertation were used in line with its aims.	* The inappropriate methods and techniques were used in the dissertation. * The data collection instrument which was not appropriate for the dissertation topic and method was used.	* The relative appropriate methods and techniques were used in the dissertation. * The data collection instrument which was relatively appropriate for the dissertation topic and method was used.	* The appropriate methods and techniques were used in the dissertation. * The data collection instrument which was appropriate for the dissertation topic and method was used.
	7. The dissertation provided new and authentic contributions to the current literature or applications.	* The dissertation results did not provide new and authentic contributions to the current literature or applications. * The results of the similar studies were repeated.	* The dissertation results provided new and but relatively authentic contributions to the current literature or applications. * The results of the similar studies were partly extended.	* The dissertation results provided new and authentic contributions to the current literature or applications. * The scope of the literature or applications was extended and different perspectives were provided.

Continued on following page

Table 8. Continued

Dimensions	Item	No	Partial	Yes
	8. The dissertation provided new and authentic suggestions to the current literature or applications.	* New and authentic suggestions were not provided to the current literature or applications based on the dissertation findings. * The suggestions made in the similar studies were provided. * The suggestions which were not based on the dissertation findings were developed.	* New and but relatively authentic suggestions were provided to the current literature or applications based on the dissertation findings. * The suggestions provided in similar studies were partly repeated.	* New and authentic suggestions were provided to the current literature or applications based on the dissertation findings. * The scope of the suggestions made in other studies was extended. * All the suggestions provided in the dissertation were consistent with its findings.
	9. The dissertation has publication potential.	* The dissertation does not have potential to be published in moderate or high impact indexed journals*. *ESCI, SSCI, AHCI, SCI, SCI Expanded, ERIC, SCOPUS indexed journals.	*The dissertation has potential to be published in moderate impact indexed journals*. *ESCI, SCOPUS, ERIC indexed journals.	* The dissertation has potential to be published in high impact indexed journals*. *SSCI, AHCI, SCI, SCI Expanded indexed journals.
CONTRIBUTIONS OF RESEARCH	10. The dissertation makes contributions to one of economic, social and cultural fields.	*Any contributions were not made to economic, social and cultural fields.	* A new but relatively authentic approach was contributed to one of economic, social and cultural fields.	A new and authentic approach was contributed to one of economic, social and cultural fields.
	11. The dissertation results make associations among different disciplines.	* The dissertation results did not make associations among different disciplines.	* The dissertation results made narrow scoped associations among different disciplines.	* The dissertation results made profound and comprehensive associations among different disciplines.

Section 3
New Perspectives on Rubrics

Moving the field forward through new perspectives, emerging contexts. Comments on limitations and potential of rubrics and points to ways forward.

Chapter 11
Knowledge of Language in Rubric Design:
A Systemic Functional Linguistics Perspective

Chahna Gonsalves

https://orcid.org/0000-0002-3516-4297

King's College London, UK

ABSTRACT

Rubrics have become popular tools for assessment and instruction in higher education. However, language choice and rubric efficacy is a topic that has been largely overlooked in the literature and teacher professional development. Composing an effective rubric—particularly for instructional and formative contexts—is a complex task that requires teachers to think about the implications of their linguistic choices for students' awareness of what and how they learn. In this chapter, the author offers a review of current research and guidance on effective rubric language. Second, this chapter uses the theory of systemic functional linguistics (SFL) to explain how SFL-informed training in rubric design can foreground language considerations to enhance teachers' capacities in effective rubric design. Overall, this chapter demonstrates that developing teachers' knowledge about language and in turn their academic and assessment literacy, is key to developing both types of literacy in students.

INTRODUCTION

Rubrics are possibly the most common means of communicating assessment expectations and criteria to students (Balan & Jönsson, 2018). The term "rubric" is often used to refer to any grading criteria. In this chapter, the term rubric refers more narrowly to a specific assessment tool presented as a matrix, which provides scaled levels of achievement for a set of assessment criteria with descriptions of various levels of the quality of performance (Allen & Tanner, 2006). This type of rubric is known as an analytic rubric (Dawson, 2017). Rubrics are used across all levels of teaching for both summative and formative

purposes (Panadero & Jönsson, 2020) and are considered valuable assessment tools and instructional tools (Jönsson & Panadero, 2017; Reddy & Andrade, 2010). By making the purpose of the task, the criteria, and performance expectations more explicit, rubrics enhance the transparency of assessment (Panadero & Jönsson, 2013), and develop students' abilities in self-assessment and evaluative judgement (Reddy & Andrade, 2010). Rubrics therefore play a key role in developing students' understanding of the process of assessment, its purpose, how it fits into their learning trajectory, and in developing their ability to judge their own work and identify means of improvement—known as assessment literacy (Chan & Luo, 2021b). Rubrics have found widespread support amongst teachers because they are a suitable vehicle for coordinating grading and comments between assessors, a time-efficient means of grading, and a sustainable and useful platform for providing feedback (Chan & Luo, 2021a).

Despite these benefits, rubrics' ability to enhance transparency of assessment has been questioned. It has been argued that teachers often take transparency for granted as students have limited understanding of rubric language, even describing it as a confusing, and thus perceive rubrics to be less helpful in clarifying aspects of assessment than teachers do (Bell et al., 2013; Fang & Wang, 2011; Li & Lindsey, 2015). Rubrics may not guarantee transparency because criteria remain opaque and valued knowledge remains implicit (Tierney & Simon, 2004). Therefore, rubrics may condition students to comply with stated criteria and standards without developing their autonomy and skills in evaluative judgement (Torrance, 2007). While the factors contributing to the efficacy of rubrics have been a topic of much research (see Panadero and Jönsson (2013, 2020) for a review), the accessibility of rubric language has emerged as a key consideration in the utility of rubrics as assessment and instructional tools (Andrade, 2001; Li & Lindsey, 2015). Empirical studies concur that rubrics can provide numerous benefits to students and staff in both assessment and supporting teaching and learning, provided they use appropriate language (Brookhart & Chen, 2015; Panadero & Jönsson, 2013, 2020). However, teachers often receive limited training and support in language and rubric development. Developing effective rubrics for assessment and instruction purposes requires that teachers understand pedagogy and assessment in their subject as well as the impact of their language choices for the users and utility of the rubric. Therefore, it is important that we support the development of teachers' knowledge of language via training and professional development.

This chapter pursues two objectives. First, an overview of current research and guidance on effective rubric language is offered. This guidance is grouped into four themes: articulating explicit criteria, pitfalls of evaluative language, precise definitions of quality and pitfalls in the pursuit of objectivity. In each theme the implications for rubric design and effectiveness are explored. Second, this chapter uses the theory of systemic functional linguistics (SFL) to explain how SFL-informed training in rubric design can foreground language considerations to enhance teachers' capacities in effective rubric design. Overall, this chapter demonstrates that developing teachers' knowledge about language and their academic and assessment literacy is key to developing both types of literacy in students. SFL offers an actionable means of achieving these outcomes. Rubric developers might take a more informed approach to word choice in the rubrics they construct, and institutional policy makers might afford language greater prominence in rubric design training and guidance.

BACKGROUND

Rubrics are often used as instruction and assessment tools in higher education (Dawson, 2017; Popham, 1997). Internationally, a changing assessment context and increasing accountability and surveillance has led to the rise of an audit culture (Edwards, 2020). Teachers are commonly required to share assessments and assessment tools such as rubrics with various stakeholders for quality assurance purposes. Simultaneously, institutional drives to move away from norm referenced tasks and improve the transparency of assessment standards and student attainment relative to learning outcomes have fueled the adoption of rubrics (Sadler, 2005) to the extent that many universities encourage or even mandate their use (Dawson, 2017). These factors have driven a need for educators to improve their assessment knowledge and abilities to design assessment tools that effectively assess and support learning. A growing body of research supports the benefits of rubrics to students and teachers (Brookhart, 2018; Jönsson & Svingby, 2007; Panadero & Jönsson, 2020; Reddy & Andrade, 2010). For example, teachers may use rubrics to provide detailed guidance on how, and on what bases, judgements about the quality of student performance will be made and how appropriate grades will be assigned (Sadler, 2005). Therefore, rubrics may help the teacher to meet their responsibility of enhancing transparency in outlining criteria openly and explicitly for the student. Transparency of criteria and assessment requirements can positively influence student learning, self-evaluation, self-regulation and improve performance. From a teacher perspective, rubrics can contribute to quality assurance aspects of assessment like reliability (Jönsson & Svingby, 2007). However, simple implementations of rubrics like just handing a rubric to students is not enough to guarantee effectiveness and teachers need training on the appropriate design, implementation, and use of rubrics (Panadero & Jönsson, 2020). Teachers require high levels of assessment literacy to effectively design and implement rubrics. Given the ubiquity of simple rubric implementations, it is important to unearth areas in which teachers need training and explore the consequences if this training is missing.

Assessment literacy is difficult to define because "the concept is still in its infancy in higher education and poorly understood" (Medland, 2019, p. 570). From a sociocultural perspective, assessment literacy is viewed as a dynamic context dependent social practice in which teachers articulate and negotiate classroom and cultural knowledges among themselves and with learners to initiate, develop and practice assessment designed to achieve the learning goals of students (Willis et al., 2013). Assessment-literate teachers must be competent to develop high quality assessments and assessment rubrics (Popham, 2009; Sadler, 1998). Yet, current training and development initiatives do not sufficiently prepare teachers for these tasks. Many members of faculty begin teaching as doctoral students, where specific training in assessment is limited, if provided at all, and opportunities for systematic professional development opportunities or guided reflection are lacking (Austin, 2002). A perception of inadequate training and preparation reverberates up the levels of academic progression, as faculty assessment leaders (Gordon & Smith, 2021) and program evaluators (Martens, 2018) also cite a dearth of formal training and preparation around assessment tools and tasks. In addition, issues of language are mostly invisible in higher education and assessment discourse (Bond, 2020; Grainger, 2021).

In the rare cases that language is considered relative to rubric design, emphasis is placed on using 'simplified' (Li & Lindsey, 2015) descriptive language to explicitly describe levels of quality or performance (e.g., Grainger & Weir, 2020), as opposed to evaluative language (e.g., excellent, poor) (e.g., Brookhart, 2018; Wiggins, 1998) which conveys a final and irreversible judgement and offers no indication for improvement. In a review of studies examining the quality and effectiveness of rubrics, Brookhart and Chen (2015, p. 343) define descriptive rubrics as "a coherent set of criteria for students' work that

includes descriptions of levels of performance quality on the criteria". In other words, a descriptive rubric is characterized by the presence of descriptions of what levels of performance might look like across a continuum of quality, however the authors offer no further discussion of how performance might be described. Without discussion of effective rubric language and its implications, such guidance may not be sufficient for most teaching staff in higher education, who are subject specialists with limited training in rubric development (Edwards, 2020), and presumably even less in linguistics.

Designing a rubric is largely a linguistic process—a complex negotiated communicative endeavor— intended to code and convey assessment expectations. Yet choosing the right words is one of the most challenging aspects of rubric design (Moni et al., 2005; Tierney & Simon, 2004). To design effective rubrics, teachers must understand how language works in a rubric to aid users in meeting the complexities of understanding, delivering, and evaluating assessment, and should also understand how academic language is learned. For teachers to design effective rubrics they require a high level of academic literacy, which refers to "the ability to communicate effectively in an academic discourse community" (Wingate, 2015, p. 6). Within a discourse community, academic literacy is a multidimensional concept which involves:

(1) an understanding of the discipline's epistemology – the ways in which subject knowledge is created and communicated, (2) an understanding of the sociological context, i.e., the status of the participants and the purpose of the interactions occurring in the community, and (3) a command of the conventions and norms that regulate these interactions. (Wingate, 2015, p. 7)

The relationship between language, assessment literacy and academic literacy is therefore complex and multifaceted. Knowledge of language can be seen as a subset of academic literacy since it involves using language to communicate in academic settings. In this case, teachers need to understand the forms and functions of language that facilitate the design and use of assessment rubrics that align with their curriculum goals, learning outcomes and community members' needs. Students need language knowledge, assessment- and academic literacy to understand and respond to assessment tasks and feedback. Knowledge of language and assessment literacy can be seen as facilitators of academic literacy since they help teachers to monitor and improve their proficiency and skills in academic and assessment contexts. For example, language knowledge and assessment literacy can help teachers to express learning and assessment goals and provide effective instruction and feedback that support learning and learner's development of academic literacy.

For students, knowledge of language can facilitate students in identifying their strengths and weaknesses in assessment through better understanding of the goals and expectations, whilst providing a common language in which to articulate goals and appropriate strategies for learning. Knowledge of language can also be seen as a product of assessment literacy and academic literacy, as it reflects the level of competence and confidence that teachers and learners have in using language for assessment purposes. Therefore, improving teachers' understanding of language and its effects expands their skills in developing and using quality rubrics, thus improving their academic and assessment literacy (Koh, 2011) as well as the academic literacy development of their students (Gebhard et al., 2014). It is also imperative that teachers—both in service and pre-service—develop this knowledge so that they can use rubrics effectively in instruction and assessment.

RUBRIC LANGUAGE AND GUIDANCE: A REVIEW

Rubrics are assessment tools that articulate the expectations of assessed work by listing criteria and for the work and describing levels of performance across a continuum of quality (Brookhart, 2018). A rubric identifies for the assessor and the assessee what mastery of the goals of the assignment look like and what to look for when assessing the work (Andrade, this volume). The following section reviews existing research and guidance on effective rubric language. Guidance relevant to two broad themes is discussed: articulating explicit criteria and pitfalls in the pursuit of objectivity.

Articulating Explicit Criteria

Articulating criteria and instances of performance presents several issues for rubric developers. First, the need to summarize the qualities of the work at various levels in a rubric, without being too rigid, can lead to descriptors that are often vague and ambiguous (Li & Lindsey, 2015). Ambiguous language impedes users' accurate and consistent interpretation of the criteria. Criteria should clearly identify the qualities of the work that demonstrate that the student has met the target learning goals, whereas descriptors should outline what these qualities look like and how they can be measured (Grainger & Weir, 2020). For example, one criterion in a rubric for a marketing project may focus on content: 'the extent to which the essay demonstrates knowledge of the marketing concepts, theories, and models relevant to the topic'. A precise descriptor for this criterion might say 'the essay shows comprehensive knowledge of the marketing concepts, theories, and models relevant to the topic through accurate selection and application in isolation and in combination.' A vague descriptor might state 'the essay shows good knowledge and understanding the marketing concepts, theories, and models relevant to the topic'.

Pitfalls of Evaluative Language

Across many curricula, the learning outcomes, and goals of assessment center on core content that all students should learn, skills or abilities to be developed and knowledge requirements. Core content is often described through nouns which provide insight into what students are expected to engage with during a course of study and/or assessment. Examples include 'analysis of data', 'production of a marketing plan', etc. However, Tan (2020, p. 30) illustrates how nouns and adjectives serve different functions in rubric criteria and performance descriptions and offer a window into rubric developers' expectations.

If we were to consider aesthetics, our expectations can be described as wanting something "beautiful," or looking for "beauty." Yet, there is a difference between the two. The word beauty, as a noun, depicts one of the many qualities we may be looking for, but not the specific level of beauty. When we look for something beautiful (rather than plain or gorgeous), the adjective is more akin to a standard. Thus, when we use a combination of adjectives or nouns to convey our academic expectations, there is a risk that we may emphasise standards or criteria, depending on whether we use adjectives or nouns. (Tan, 2020, p. 30)

The language used in assessment rubrics should be descriptive, and not evaluative, such that the language helps students understand how they have met the learning goals and where they are going rather than simply indicating whether their performance was good or poor. Descriptive phrases are those that describe the quality or characteristics of the student's work, without making a judgment about whether it

is good or bad. Evaluative phrases are those that assign a value or grade to the student's work, based on some criteria or standards. For example, suppose you have a criterion for assessing the student's ability to apply marketing concepts to a real-world scenario. A descriptive phrase for this criterion might be 'the student demonstrates a clear and accurate understanding of the relevant marketing concepts and applies them appropriately to the scenario'. In contrast, an evaluative phrase for this criterion might read 'the student's application of marketing concepts is excellent and shows a high level of mastery'. In this example, the descriptive phrase focuses on what the student did and how well they did it, while the evaluative phrase assigns a label or rating to the student's performance. According to Brookhart (2018), descriptive rubrics are generally expective to be most conducive to learning. Yet, writing descriptive as opposed to evaluative performance descriptors presents a challenge for even the most experienced rubric designers. As a result, many examples contain a mix of the two (Brookhart, 2018). Table 1 provides examples of descriptive versus evaluative descriptors.

Table 1. Descriptive versus evaluative performance descriptors in rubrics

Descriptive	Evaluative
The student identifies and explains the main marketing objectives and strategies of the chosen company.	The student's analysis of the company's marketing objectives and strategies is thorough and insightful.
The student uses relevant and credible sources to support their arguments and cites them correctly using APA style.	The student's use of sources is excellent and demonstrates a high level of academic integrity and skill.
The student demonstrates creativity and originality in their approach to the marketing problem and solution.	The student's marketing solution is innovative and unique, standing out from the rest of the class.
The student organizes their report in a clear and logical manner, with an introduction, body, conclusion, and references.	The student's report is well-structured and follows the required format.
The student communicates their ideas effectively and professionally, using appropriate language, tone, and grammar.	The student's writing style is impressive and error-free, showing a strong command of the English language.

Through working with the core content, students can be guided to reach the knowledge requirements, which are often defined through verbs with adjectives of adverbs for qualifiers or achievement against the standards. Students find the use of evaluative adjectives such as 'outstanding', 'excellent', 'good', 'fair', or 'poor' (whether they are appreciative or derogatory) misleading and confusing; they perceive them to be devoid of meaning and dismiss them in self-evaluation because these words invoke subjective interpretation (Wang, 2017). Li and Lindsey (2015) examine teachers' and students' perceptions of rubric used on an undergraduate writing course and found discrepancies in both parties' understanding of evaluative language used in the rubric. They conclude that rubrics did not always clarify assessment expectations for students as effectively as teachers assumed they would, due to differences in interpretation of evaluative language. One way to avoid these differences in interpretation is to describe the qualities or characteristics of the work that demonstrate achievement of the learning goals. For example, instead of stating instead of saying "The student's report is good', a descriptive rubric might say 'the students report contains evidence, analysis, and recommendations.' Descriptive language, as opposed to evaluative language, helps the student to gauge the extent of their learning, to direct progression on their learning journey, and to specify learning goals (Brookhart, 2018). For example, a descriptor stating 'the research uses a balance of verbal facts/numerical data which are examined and discussed as component parts of an argument' denotes for a student what information and treatment might evidence 'analysis'.

In contrast, indicating work 'well analysed' at a particular level of performance is evaluative and tells the student little about what good analysis might look like.

Second, ambiguity may be used to set the tone of rubric and covertly indicate bandwidth of professional judgement for assessors. Interpreting rubric criteria and descriptors entails subjectivity (Sadler, 2011). Some teacher trainers may suggest the inclusion of adjectives and adverbs to "soften" rubric language, which can increase the assessors' subjectivity (Holmes & Oakleaf, 2013). In assessment rubrics, adverbs like 'very' more often than adjectives, are the qualifiers that refer to proficiency. Kohn (2006, p. 12-13), who takes a position of critique against the general idea of rubrics points out that:

Rubrics are, above all, a tool to promote standardization, to turn teachers into grading machines or at least allow them to pretend that what they are doing is exact and objective… [However, they are criticised because] they can never deliver the promised precision; judgments ultimately turn on adjectives that are murky and end up being left to the teacher's discretion.

Gipps (1999, p. 370) further highlights that judgement of performance in particular "is construed according to the perspectives and values of the assessor". Experienced assessors have a tacit understanding of what constitutes high-quality student work and can recognise high quality when they see it (Lea & Street, 1998; Sadler, 2013). Grainger et al. (2008) find that using this tacit knowledge, experienced assessors are able to accurately interpret descriptions of student behaviour that are often ambiguous in order to differentiate between standards or levels of performance. However, rubric designers should not assume assessors are experienced. Rubrics should be written a language effective for graduate teaching assistants and experienced assessors alike. More precise use of language facilitates universal translation and interpretation of the rubric in the way the rubric developer intended (Jönsson & Svingby, 2007; Libarkin, 2008). To improve precision, some rubric designers may opt for longer performance descriptors. Some research does in fact indicate that rubrics with more detailed descriptors significantly increase inter-rater reliability and improve raters' ability to identify different aspects of the work (e.g., Knoch, 2009), although longer descriptors do not necessarily result in greater effectiveness for students (Covill, 2012; Li & Lindsey, 2015).

Precise Definitions of Quality

Precision is increased in the rubric when developers use qualitative instead of quantitative terms (Wiggins, 1998). Qualitative language such as 'consults and interprets appropriate sources' describes the quality of work, whereas quantitative terms such as 'has three sources' can lead to a laundry list of points of requirements (Brookhart & Chen, 2015). Statements related to the number of times a student has demonstrated a skill, or similarly, the number of times an error is observed, as an indicator of the students' level of mastery lead only to an illusion of precision while overlooking how well the student has performed in the work (Holmes & Oakleaf, 2013). Related are frequency terms like 'always', 'sometimes', 'rarely', 'never' which are commonly used as rating scales (Brookhart, 2018) and find their way into performance decsriptors, but do not describe performance. When designing the rubric, the teacher must ask whether the number of times, or the frequency with which a skill or error manifests can truly establish where a student is in their learning. Does demonstrating a skill once rather than three times reveal anything about the depth of students' mastery of that skill? If so, what does developmental feedback look like for that

student, and can an improvement be gleaned from comparing their performance with the next level of performance descriptor in the rubric?

Articulating the gradation of performance can be difficult because words may not be able to capture the nuance and richness of performance (Sadler, 2011). Therefore, being completely explicit in a rubric may be more difficult than it sounds. Bearman and Ajjawi (2018) suggest that true precision and explicitness may not even be possible. However, the exercise of articulating criteria precisely forces rubric designers to identify what they consider to be relevant and important as demonstrations of learning. Defining performance as precisely as possible allows for reflection, discussion within a community of practice, critique, and later modification. Developing teachers' knowledge and mastery of rubric language, and in turn, their academic and assessment literacy, is key in developing their ability to produce effective rubrics.

Pitfalls in the Pursuit of Objectivity

Guidance on rubric language commonly suggests that descriptors should be written in objective language as far as possible to facilitate objective assessment, which is not influenced by feelings, opinions, or implicit judgement in representing facts of the assessed performance (Chan & Ho, 2019; O'Donnell et al., 2011). Objectivity is enhanced when developers refrain from using adverbial phrases such as 'sound understanding of' and 'comprehensive use of' or adverbs like 'very' and 'little' (Grainger, 2021). However, this apparently simple advice overlooks several considerations.

First, difficulties in generating descriptors that meaningfully differentiate between performance levels can lead to the use of evaluative and comparative language, such as 'provides adequate', 'sound', or 'excellent argument' (O'Donnell et al., 2011). The use of adverbs may mask the developer's attempt to capture some of the tacit knowledge that simply cannot be codified or expressed (Bearman & Ajjawi, 2018). Another example of this problem is descriptors stating the student 'uses some eye contact' and 'student uses eye contact effectively' without specifying what 'some' or 'effectively' mean (Oakleaf, 2009). Attempts to capture this knowledge in the rubric necessarily lead to simplification and lose some sense of the knowledge we translate into words, thus introducing some subjectivity. Some guidance on writing effective rubrics advises rubric designers to articulate the gradation of performance quality by describing the best and worst levels of quality and then filling in the levels in between based on their knowledge of common problems (e.g., Montgomery, 2000). However, anchoring the generation of descriptors around common problems that may occur invariably leads to the use of negative language (e.g., 'does not identify some parts of the problem', 'identifies the problem but does not propose a solution'). Descriptors of lower levels of performance should not use unduly negative language (Oakleaf, 2009). Negative descriptions may demotivate students. Besides clarifying the negative qualities of work evidenced at the lowest or failing level, rubric wording should focus on what the student demonstrated they can do (Grainger & Weir, 2020).

Another problem with negative performance decsriptors at the lower levels of a rubric is that they do not indicate to students who are clearly struggling how to improve, or what their next steps should be. If rubrics can and should support learning, the descriptions at all levels should provide instruction and guidance, not evaluative put-downs. Positive qualities, which are precisely distinguished at the various levels of performance, are also potentially easier for students to evidence and for assessors to identify concretely at each incremental level of success. Rubric developers might consider describing effective performance at each level as well as identifying problems commonly demonstrated in the work to better differentiate between performance levels in positive terms. For example, 'identifies both a problem and

solution', 'offers critical assessment of proposed solution considering perspective other than their own'. Research shows that students do not consider opposing arguments, not because they are not able to, but because they do not think of doing so (Andrade, 2001; Perkins et al., 1993). Thus, identifying common problems can help to distinguish between levels of performance more precisely while also enhancing the rubric's developmental value for users' academic and assessment literacy (Andrade, 2001).

Second, and compounding the first difficulty, is that rubric developers may not want to design descriptors that are overly prescriptive for students or other assessors. Therefore, they may use language of uncertainty, known as hedging. Hedging refers to words that indicate tentativeness and possibility (e.g., 'may', 'might', 'could', 'likely', etc.) (Hyland, 1996). Hedge words also appear in descriptors differentiating levels of performance with statements such as 'to some extent', 'more', or 'less'. For these terms to be interpreted accurately, the rubric designer needs to give the assessor some description of what 'to some extent' looks like in students' work. The subjectivity and diversity of interpretation that hedged statements can produce thwarts the aim of unambiguous rubric communication. Moreover, rubric developers' proclivity for hedging may preclude true evaluation or benchmarking of performance by creating a scenario where descriptors overlap, so a student's performance can be pigeonholed against multiple descriptors, or may fit into any (Oakleaf, 2009). Descriptors should convey an absolute basis for performance to reduce the need for them to be interpreted relative to each other. Sadler (2014, p. 281) specifies that qualifiers, modifiers, and hedge words should thus be avoided. In addition, Sadler points out that "A work that is 'outstanding' literally stands out from some background, either real and immediate, or recalled from memory", meaning that a judgement of "outstanding" can only be made if the performance is judged relative to other work. The same applies to extraordinary, exceptional, excellent, and superb. Assessors may not have access to the full breadth of student work to make comparative judgements, and even if they do, may not have access to the same sample as other assessors which inevitably cultivates unreliability in the evaluation processes. Moreover, students will rarely, if ever, have access to the complete body of student work to make judgements of their own their own performance relative to others.

Beyond assessing content knowledge, rubrics provide a platform for the assessment of processes, skills, competencies, and quality of performance or product. Processes (e.g., 'details the methods and or steps of data analysis' or 'systematically explains analytical reasoning'), competencies (e.g., 'applies basic negotiation principles to diffuse conflict and strengthen relationships between groupmates') and skills such as creativity of analysis (e.g., 'combines analysis techniques for nuanced interpretation') must be articulated to permit their observation and measurement. It may not be possible to express measures of process, skills, competencies, performance, and product in a single rubric. In a study of how creativity can be assessed, Lindström (2006) used separate rubrics to detail product versus process criteria. Similarly, Cole et al. (2018) demonstrated how assessment of STEM students' learning could be facilitated by using two rubrics: a product rubric to assess students' written or submitted work, and an interaction rubric to assess students' behaviour, processes, and workplace skills during active learning. Rubric designers must be especially clear on the indicators that evidence processes, competencies or skills may be demonstrated in student work. How rubric designers convey the complex requirements of their assessment via their language is a key aspect of their assessment literacy which greatly impacts the assessment literacy of their assessor team and their students.

However, gains in clarity may come at the cost of rubrics' effectiveness for the student. Enhanced clarity may facilitate student success and foster learning (Andrade, 2000; Jönsson & Panadero, 2017), but over-specification in rubrics can diminish student agency and learning, leading to 'criteria compli-

ance' (Torrance, 2007). Studies show that greater elaboration of criteria and performance descriptors add little value in clarifying criteria for students (Huang, 2012), and instead can accentuate the subjectivity built into the rubric (Turley & Gallagher, 2008). To promote learning in the short-term, rubric language should nurture students' semantic and contextual understanding and develop their skill in evaluative judgement. In the long-term, familiarisation and fluency with rubric language promotes development of their academic and assessment literacy. Therefore, inducting assessors and students to the language to establish understanding and achieve consensus of interpretation is fundamental (Sadler, 2011). Rubric developers' knowledge of the language they use necessarily plays a role in their ability to streamline interpretation. In addition, achieving objectivity of some descriptors might obscure or diminish the focus on others (Bearman & Ajjawi, 2018). Recognising how the language used emphasizes or downplays aspects of the rubric allows the developer to strike an intended balance and identify where understanding of fundamental concepts or vocabulary and socialisation to the discourse may need to be supported in other ways (Walker & Hobson, 2014).

In addition to evaluation or assessment terms, a rubric may feature subject-specific content words. Content words are terms that denote concepts, principles, technical terms, or jargon relevant to the assessment, subject, and discourse. For instance, using the example of marketing, students must understand that marketing means 'to market' ideas, people, places, products, or services, and also refers to the activities and process involved in transferring such offerings from producer to buyer or user such a that a mutually beneficial exchange takes place (Brennan & Vos, 2021). Content words featured in a rubric are likely to be critical to students' grasp of the subject and completion of the assignment. Therefore, students must understand this language deeply. However, these words may not be familiar to students who have not studied this subject before. Content words become even more problematic in a rubric when students' ability to access or understand these words depends on their understanding of other vocabulary.

Further complicating the issue, supporting words may have homophones or different meanings across subjects or disciplines (Calderon & Slakk, 2018). A rubric that states 'the student demonstrates an understanding of the ethical implications of the chosen research topic and the informed consent process for the participants' requires the student to be familiar with the varous ways the word 'ethics' may be used in different subjects, to understand what ethical implications are and how they relate to research. Therefore, it is important that students have access to definitions, explanations, or examples of these content words, either in the rubric itself or in other sources such as handouts explaining terms and glossaries (Cox et al., 2014). However, glossary definitions per se are not explicit (Richards & Pilcher, 2014). Glossaries should be adapted to the assessment or subject-specific discourse, and shared understanding of glossary words might be reached using teacher-led explanation and dialogue (Hawe et al., 2021; O'Donovan et al., 2008). The semantic meaning of rubric words will differ across subjects and contexts (Cox et al., 2014). Therefore, rubric developers might consider when delineation of meaning is important and the circumstances under which multiple, or alternative, meanings might arise.

A SYSTEMIC FUNCTIONAL VIEW OF RUBRIC LANGUAGE

SFL (Halliday & Kress, 1976; Halliday & Matthiessen, 2014) considers language to be a system of subsystems, where its function and meaning are derived from its use in a social context. According to Halliday and Matthiessen (2014), language serves three main functions, which they call metafunctions. These are: (1) ideational, which serves to organise our experience in the world; (2) interpersonal, which

serves to express interactions; and (3) textual, in which linguistic units are contextualized and organised as discourse.

Whereas linguistic theories tend to describe language as a set of rules that are characteristic of a context, SFL describes language as systems of choices that users can draw upon based on the meaning they wish to impart. Choice relates to the selection of language features. The notion of choice is a fundamental concept in the theory of SFL, as it is through choice that meaning is created and expressed (Halliday, 2013). With regards to language features, an SFL lens highlights the importance of what is chosen and what is not but could have been (Fontaine, 2013). The impression given in SFL literature about choice "implies *more* than the mere availaibility of features in an inventory but *less* than a deliberate, purposeful communicative act" (Bache, 2013, p. 73). Consideration of the purposeful meaningful communicative goals we try to achieve with the aid of assessment rubrics is important because langauge choice closely reflects and impacts what we choose to mean (Halliday, 2013). Moreover, by treating these choices as a motivated means to a communicative end, allows for greater achnowledgement that choosing to mean in an assessment rubric is often guided, but can be modified, by design.

SFL also conceives language as a social process, connected to a society or community and its culture. The general rhetorical practices of a culture, called genres (Martin & Rose, 2008), are characterized by discourse organization and by obligatory and optional language features (Brisk & Zisselsberger, 2011). Within a genre, language use serves a social function, where its construction and interpretation of meaning are shaped by the social, cultural, and situational context. In this functional view of language, language use varies depending on the context. That is, language is used differently in different situations and for different different audiences (e.g., novice students, expert markers) subject areas (e.g., marketing, history) and purposes (e.g., teaching, self assessment, expert marking). In SFL, language use is learned in and from context. Therefore, to fulfil the requirements of the genre and to design a rubric that effectively achieves it communicative goals depends on one's knowledge of language, ability to use it in context and understanding of the audience.

Genres can be viewed through three lenses: field, which relates to how the content of the text is expressed; tenor, which relates to the relationship between the interactants; and mode, which relates to the way that language is delivered (e.g., written or spoken, formal or informal, etc.). Because this chapter's focus is written rubrics, mode is not discussed further. These aspects of genre are construed through and correspond to the three planes of meaning in SFL, called metafunctions, which are ideational, interpersonal, and textual (Halliday & Matthiessen, 2014). From the SFL perspective, genres are typified by texts with specific structures and purposes to fulfil communicative and/or social goals (Martin & Rose, 2008). Question and answer texts, essays, and rubrics are all genres of text produced which are related to the university assessment situation. Systemic variation of field, tenor, and mode leads to distinguishable variation the language, relative to its usage in a particular situation such as academic register (Halliday & Hasan, 1989). Table 2 shows the relationship between the three SFL metafunctions and language choices, which are relevant to the rubric genre.

Table 2. SFL metafunctions and language choices in rubrics

Metafunction	Genre aspect	Unit of analysis	Description
Ideational	Field	Content words	Subject-specific concepts
		Participants (nouns)	Words used to describe a class of entities actions, qualities, states of existence, or ideas.
		Processes (verbs)	Words used to convey an action, occurrence, method, or state of being.
		Attributes (adjectives)	Words naming the attributes of a noun.
		Quality (adverbs)	Words modifying the participants (nouns), processes (verbs), or attributes (adjectives).
Interpersonal	Tenor	Positive versus negative words	Words describing describe something good or desirable and express approval, possibilities, and potential, versus bad, undesirable, disapproval, impossibility, and failure.
		Modal verbs/adverbs	Words used to indicate likelihood, ability, permission, request, capacity, suggestions, order, obligation, or advice, (e.g., can/could, may/might, will/would, shall/should, must, etc.).
		Evaluation-laden lexical choices	Words which express the subjectivity of the marker (e.g., excellent, good, adequate, inadequate, poor).
		Familiarity and formality	Academic tone denoting status and relationships as well as the intended user of the rubric (e.g., assessors versus students); formality of language used to facilitate the user in evaluation.

Adapted from Aguirre-Muñoz et al. (2009) and Troyan et al. (2019).

As an applied model of linguistics, SFL is designed to be a strategic tool and a guide to action, as well as a means of responding to everyday language-related issues in a variety of social, professional, and academic contexts. The students that enter university education are increasingly diverse in terms of language, cultural background, ethnicity, disability status, learning needs, and educational experience (McDuff et al., 2020). For some students, rubrics can be helpful in translating the requirements and expectations of assessment, and for many, an effective rubric can help to level the playing field for engagement with the assessment and subsequent attainment (Smith et al., 2013). However, students often have very different understandings, or little understanding, of the academic terms we use. Stevens and Levi (2013, p. 42) point out that:

…we may be startled to discover that many students think 'introduction' and 'conclusion' are synonyms for 'beginning' and 'end' or that 'critical thinking' means criticizing something… some students may assume that 'analysis' refers only to situations in which numbers are involved or to the analysis contained in secondary sources. The revelation that in an academic paper, for example, 'analysis' most often means their own conclusions informed by data can be startling to them.

Although rubrics may be beneficial to students throughout the progression of their course, they may be most influential in the development of assessment literacy and socialisation to disciplinary discourse in the formative years of learning. Gebhard (2010) and Schleppegrell (2004) illustrate the utility of SFL-based pedagogies in enhancing students' familiarity with the linguistic options relevant for defining concepts, making an argument, describing process, and narrating an event. Gebhard (2010) offers the example of an interaction with school-aged student during an activity designed to improve students' understanding

of scientific method and the language for objectively describing the findings of an experiment. Gebhard corrected the student's use of the first-person to describe findings and explained that scientists typically report their studies in objective and author evacuated language. The student expressed her reluctance to speak like "a geek". However, her response revealed that she understood that linguistic choices construct and conveys ideas, reflect modes of interaction (familiar or formal), and convey aspects of identity. Schleppegrell (2004) highlights that the content and medium of teaching as well as the demonstration of learning are all dependent on academic language. Teachers often assume students' fluency with academic language (Stevens & Levi, 2013), yet facilitating students' understanding of the language used in a rubric is an important element of effective rubric design and assessment literacy (Smith et al., 2013). Therefore, SFL is a valuable lens for language-based approaches in instruction as users learn language, learn through language, and learn about language (Halliday, 1993). SFL-based pedagogies focus on expanding users' awareness of the linguistic options available to them to engage with assessment and broaden their ability to use language more expertly to achieve assessment tasks. This statement applies whether we take the student or the teacher as the user.

A concerning finding in research is that teachers are often unable to clearly articulate what they mean by terms they regularly use in assessment. For example, Lea and Street (1998) found that even though assessors are able to identify successful and unsuccessful work, they were not able to describe what constitutes a 'well-argued' or 'well-structured' piece, or how a particular piece of writing 'lacked' structure. Academics also had difficulty explaining terms such as 'critical analysis' and 'evaluate'. Lea and Street quote one lecturer who stated "'I know a good essay when I see it, but I cannot describe how to write it" (p. 163). Terms like 'critical thinking' feature prominently in assessment and rubrics, but assessors' opinions and interpretations of what constitutes good or poor quality critical thinking differ (Bloxham et al., 2011). Effective communication of criteria and expectations depends on shared definitions, assumptions, and understanding. An SFL perspective views language as a set of resources, rather than a set of rules. This "makes it possible to consider the appropriateness or inappropriateness of language choices in a given a context of use" (Gibbons, 2003, p. 250-251). Therefore, an SFL approach to rubric design serves to encourage rubric designers to clarify these definitions, become aware of their language beliefs and assumptions, and address dissonance in understanding, iteratively, until it is certain that all rubric users are speaking the same language. In more practical terms, this means that An SFL perspective to rubric design not only raises the awareness for teachers and assessors about the academic language demands but equips them to anticipate student need and be prepared for discussion and teachable moments. An SFL approach may also reveal which aspects of language for which understanding may need facilitation, allowing a teacher to focus on those. SFL is widely used in Australia, and increasingly in the United States as a response to teacher preparation programmes that often lack sufficient focus on the linguistic structures that characterize academic language (Aguirre-Muñoz et al., 2009, p. 298).

IMPLICATIONS

In this chapter, an SFL approach to rubric design has been introduced and its value for improving teachers' and students' academic and assessment literacy has been discussed. It also provided an overview of extant research and guidance on effective rubric language with a discussion of the implications for rubric design and effectiveness. Developing teachers' understanding of the language they use in their rubrics by drawing on SFL scholarship could enhance their assessment literacy and their ability to de-

sign effective rubrics. Rubrics designed with mindfulness of linguistic choices and the consequences of those choices are likely to be of greater utility to their users. When used as a learning support tool for students, rubrics designed with greater consideration of language may increase students' awareness of what and how they learn.

Developing rubrics shapes how teachers think and teach, not only because rubrics provide the framework in which they situate their expectations, but also because they force teachers to articulate and codify those expectations. This activity is perhaps most influential to the development of teacher's assessment literacy in the formative years of teaching. However, as their socialisation into the language of their discipline evolves, so will the way they use, understand, and assess against that language. Teacher trainers have an essential function in guiding teachers' practice. Without sufficient focus on the non-technical aspects of rubric development, the benefits of rubrics may be called into question.

Teacher education, colleges, and universities seeking to improve faculties' capacities in designing effective rubrics can take several approaches to introducing SFL. First, institutions favouring a short-term approach might offer group workshops designed to provide teachers (i.e., rubric developers) with a foundational understanding of how language works in the rubrics they routinely provide to their students, informed by SFL scholarship. Such workshops might use SFL resources to introduce the metafunctions of language and link these to rubric language using exemplars or participants' own rubrics. In SFL-based professional development, teachers are typically introduced to functional terms for noticing and naming the features of genre and or/register, which develops their ability to use that language to analyse a text (Accurso & Gebhard, 2021). Through analysing rubrics, teachers develop more critical understandings of the text in context through more conscious awareness of the meaning-making intents and purposes of the text (de Oliveira et al., 2021; Schleppegrell, 2004). This would facilitate teachers' active learning of effective rubric language as well as reflection and discussion of the impact of language on their rubrics' outcomes. Galguera (2011) suggests that it may be through providing teachers with opportunities to examine specific functions of language in academic contexts and experience ways in which language is used to represent knowledge in rubrics and interpersonal dynamics encoded in language, that they begin to construct deep understandings of language. This type of professional development can help teachers to explore the meaning and the design of meanings within specific disciplinary discourse communities which fosters their ability to guide other assessors and students to do the same (Accurso & Gebhard, 2021). With deeper and more critical understanding of the impact of word choice, teachers would be better placed to design rubrics that explicitly focus on and support the development of pedagogic knowledge and disciplinary and assessment literacy practices.

Empirical studies reveal that SFL training in workshops has positive impacts on teachers' awareness of language and their capacity to develop rubrics that serve as effective instructional and assessment tools. Berg and Huang (2015) investigated how integrating an SFL approach to skill and instructional practice development in teacher training impacted their instruction of language in disciplinary content areas. Their study demonstrated that the SFL approach improved teacher candidates' linguistic sensitivity and enhanced their ability to design instruction within their subject with an explicit focus on language and literacy development. Similarly, Brisk and Zisselsberger (2011) demonstrated that participating in a teacher development program that introduced SFL theory and strategies to teach specific genres of writing improved teachers' confidence to teach those genres of writing, and to develop and review their lessons with language features in mind. In a study of 21 teachers, Aguirre-Muñoz et al. (2009) examined the instructional impact of a one-week SFL-informed teacher development program designed to improve teachers' understanding of the linguistic features of academic language in their disciplines.

Using evaluations of student work prior to and after the SFL training, the findings revealed that teachers' understanding of the linguistic features of their subject area had improved and they were better able to evaluate student work in a way that informed further instructional needs. Taken together, these findings suggest that teacher training need only introduce the concept and general principles of SFL, drawing on the literature and scholarship available, to give teachers the tools they need to think more critically about rubric language.

A concern with the complexity of SFL theory is briefly mentioned in the literature; thus, teacher trainers should exercise caution. SFL theory is useful for those who learn it because it is well suited to use outside the classroom and can be easily applied to everyday life, yet those who teach it may find it to be "a veritable maze, very messy and complex" (Bourke, 2005, p. 93). However, Macken-Horarik (2008, p. 47) highlights that SFL-informed training can be useful provided teachers have a level of SFL metalanguage that is good enough "to think with", but that doesn't require them to become theoretical linguists. Awareness of this metalanguage and some familiarity with it provides a means to foreground language considerations in rubric design and instruction and connect this knowledge in a meaningful way to the social purposes for which the language is being used. Such an approach positions SFL as just one tool in a box of larger efforts to aid teachers in developing knowledge of language, to critically assess students' emergent academic literacy practices and to implement responsive pedagogical practices (Gebhard et al., 2008). It may also be crucial that teacher trainers are frequently reminded that rubric language, which is genre-specific, cannot be presented as a fixed set of rules due to the important role context plays in guiding word choice and meaning for the rubric's audience and purpose (Brisk & Zisselsberger, 2011).

Once teachers are aware of SFL's concepts, another approach to developing faculty's abilities to design rubrics to enhance academic and assessment literacy is for faculty to act as peer reviewers, giving each other feedback on rubrics. Faculty review of rubrics can lay fertile ground for sharing best practice in rubric language and promotes continued active learning from a range of examples and perspectives, due to faculty members' dual roles as rubric developers and members of marking teams. Because faculty will approach rubric design with varying levels of socialisation into the language of the discipline and experience, it is important for institutions to facilitate training and professional development opportunities that support growth at the novice, intermediate, and advanced levels. While novice teachers will need guided opportunities to develop rubrics and reflect on the effectiveness of rubric language, they also need opportunities to develop their pedagogical and assessment approaches to know how the criteria might be realised. More experienced teachers are likely to have more awareness of how criteria can be realised, leading them to focus on the nuances in communicating levels of performance in a way that generates consistent interpretations among rubric users.

CONCLUSION

In sum, the rationale for improving students' awareness of what and how they learn via rubrics supports the need for equipping teachers with contemporary knowledge about the language of effective rubrics and this language's relevance in developing students' and assessors' academic and assessment literacies. This knowledge simultaneously develops the teacher's academic and assessment literacy and allows them to make intentional and informed rubric design choices. Rubric training and teacher development programs could adopt training in SFL. Training in SFL would systematically build rubric developers'

and assessors' linguistic knowledge base, allowing them to develop a nuanced understanding of the discourse and context through field and tenor. This understanding could also help to focus teaching and learning on the way language functions within the discourse, raising individuals' awareness to areas where language should be given greater visibility during instruction. It is hoped that the review offered in this chapter assists teachers by increasing their awareness of language and its implications to enhance the effectiveness of their rubrics in developing users' academic and assessment literacy. Institutions should give higher priority to considerations of rubric language for teachers who have limited pedagogical training—which leaves them underprepared for effective rubric design. The aspiration is that this discussion might prompt the inclusion of SFL concepts in training and professional development focusing on rubric design.

REFERENCES

Accurso, K., & Gebhard, M. (2021). SFL praxis in US teacher education: A critical literature review. *Language and Education*, *35*(5), 402–428. doi:10.1080/09500782.2020.1781880

Aguirre-Muñoz, Z., Park, J.-E., Amabisca, A., & Boscardin, C. K. (2009). Developing teacher capacity for serving ELLs' writing instructional needs: A case for systemic functional linguistics. *Bilingual Research Journal*, *31*(1-2), 295–322. doi:10.1080/15235880802640755

Allen, D., & Tanner, K. (2006). Rubrics: Tools for making learning goals and evaluation criteria explicit for both teachers and learners. *CBE Life Sciences Education*, *5*(3), 197–203. doi:10.1187/cbe.06-06-0168 PMID:17012210

Andrade, H. G. (2000). Using rubrics to promote thinking and learning. *Educational Leadership*, *57*(5), 13–19.

Andrade, H. G. (2001). The effects of instructional rubrics on learning to write. *Current Issues in Education (Tempe, Ariz.)*, *4*(4). http://cie.ed.asu.edu/volume4/number4/

Austin, A. E. (2002). Preparing the next generation of faculty: Graduate school as socialization to the academic career. *The Journal of Higher Education*, *73*(1), 94–122. doi:10.1080/00221546.2002.11777132

Bache, C. (2013). Grammatical choice and communicative motivation: a radical systemic approach. In *Systemic functional linguistics: Exploring choice* (pp. 72–94). Cambridge University Press. doi:10.1017/CBO9781139583077.006

Balan, A., & Jönsson, A. (2018). Increased explicitness of assessment criteria: Effects on student motivation and performance. *Frontiers in Education*, *3*, 81. doi:10.3389/feduc.2018.00081

Bearman, M., & Ajjawi, R. (2018). From "seeing through" to "seeing with": Assessment criteria and the myths of transparency. *Frontiers in Education*, *3*(96), 1–8. doi:10.3389/feduc.2018.00096

Bell, A., Mladenovic, R., & Price, M. (2013). Students' perceptions of the usefulness of marking guides, grade descriptors and annotated exemplars. *Assessment & Evaluation in Higher Education*, *38*(7), 769–788. doi:10.1080/02602938.2012.714738

Berg, M. A., & Huang, J. (2015). Improving in-service teachers' effectiveness: K-12 academic literacy for the linguistically diverse. *Functional Linguistics*, *2*(1), 1–21. doi:10.118640554-015-0017-6

Bloxham, S., Boyd, P., & Orr, S. (2011). Mark my words: The role of assessment criteria in UK higher education grading practices. *Studies in Higher Education*, *36*(6), 655–670. doi:10.1080/03075071003777716

Bond, B. (2020). *Making language visible in the university*. Multilingual Matters.

Bourke, J. (2005). The grammar we teach. *Reflections on English Language Teaching*, *4*(2), 85–97.

Brennan, R., & Vos, L. (2021). Introduction to teaching marketing. In R. Brennan & L. Vos (Eds.), *Teaching marketing*. Edward Elgar Publishing. doi:10.4337/9781789907896.00008

Brisk, M., & Zisselsberger, M. (2011). "We've let them in on a secret": Using SFL theory to improve the teaching of writing to bilingual learners. In T. Lucas (Ed.), *Teacher preparation for linguistically diverse classrooms: A resource for teacher educators* (pp. 111–126). Routledge.

Brookhart, S. M. (2018). Appropriate criteria: Key to effective rubrics. *Frontiers in Education*, *3*(22), 1–12.

Brookhart, S. M., & Chen, F. (2015). The quality and effectiveness of descriptive rubrics. *Educational Review*, *67*(3), 343–368. doi:10.1080/00131911.2014.929565

Calderon, M. E., & Slakk, S. (2018). *Teaching reading to English learners, grades 6-12: A framework for improving achievement in the content areas*. Corwin Press.

Chan, C. K. Y., & Luo, J. (2021a). Exploring teacher perceptions of different types of 'feedback practices' in higher education: Implications for teacher feedback literacy. *Assessment & Evaluation in Higher Education*, *47*(1), 61–76. doi:10.1080/02602938.2021.1888074

Chan, C. K. Y., & Luo, J. (2021b). A four-dimensional conceptual framework for student assessment literacy in holistic competency development. *Assessment & Evaluation in Higher Education*, *46*(3), 451–466. doi:10.1080/02602938.2020.1777388

Chan, Z., & Ho, S. (2019). Good and bad practices in rubrics: The perspectives of students and educators. *Assessment & Evaluation in Higher Education*, *44*(4), 533–545. doi:10.1080/02602938.2018.1522528

Cole, R., Lantz, J., Ruder, S., Reynders, G., & Stanford, C. (2018). Enhancing learning by assessing more than content knowledge. ASEE Annual Conference & Exposition, Salt Lake City, Utah, Covill, A. E. (2012). College students' use of a writing rubric: Effect on quality of writing, self-efficacy, and writing practices. *Journal of Writing Assessment*, *5*(1).

Cox, K., Imrie, B. W., & Miller, A. (2014). *Student assessment in higher education: A handbook for assessing performance*. Routledge. doi:10.4324/9781315042107

Dawson, P. (2017). Assessment rubrics: Towards clearer and more replicable design, research and practice. *Assessment & Evaluation in Higher Education*, *42*(3), 347–360. doi:10.1080/02602938.2015.1111294

de Oliveira, L., Smith, S. L., Axelrod, D., Diaz, E., & Vicentini, C. (2021). Supporting Academic Language Development for Multilingual Learners Across Content Areas Through the Identification of Textual Features. *Journal of Narrative and Language Studies*, *9*(17), 227–242.

Edwards, F. (2020). Engaging tertiary educators in the development of their assessment literacy. *Teachers and Curriculum*, *20*(1), 87–92. doi:10.15663/tandc.v20i1.345

Fang, Z., & Wang, Z. (2011). Beyond rubrics: Using functional language analysis to evaluate student writing. *Australian Journal of Language and Literacy*, *34*(2), 147–165. doi:10.1007/BF03651853

Fontaine, L. M. (2013). Introduction: Choice in contemporary systemic functional theory. In L. Fontaine, Tom Bartlett, & G. O'Grady (Eds.), Systemic functional linguistics: Exploring choice. (pp. 1-12). Cambridge University Press.

Galguera, T. (2011). Participant structures as professional learning tasks and the development of pedagogical language knowledge among preservice teachers. *Teacher Education Quarterly*, *38*(1), 85–106.

Gebhard, M. (2010). Teacher education in changing times: A systemic functional linguistics (SFL) perspective. *TESOL Quarterly*, *44*(4), 797–803. doi:10.5054/tq.2010.237335

Gebhard, M., Chen, I.-A., & Britton, L. (2014). "Miss, nominalization is a nominalization:" English language learners' use of SFL metalanguage and their literacy practices. *Linguistics and Education*, *26*, 106–125. doi:10.1016/j.linged.2014.01.003

Gebhard, M., Demers, J., & Castillo-Rosenthal, Z. (2008). Teachers as critical text analysts: L2 literacies and teachers' work in the context of high-stakes school reform. *Journal of Second Language Writing*, *17*(4), 274–291. doi:10.1016/j.jslw.2008.05.001

Gibbons, P. (2003). Mediating language learning: Teacher interactions with ESL students in a content-based classroom. *TESOL Quarterly*, *37*(2), 247–273. doi:10.2307/3588504

Gipps, C. (1999). Socio-cultural aspects of assessment. In A. Iran-Nejad & P. D. Pearson (Eds.), *Toward a new science of educational practice. Review of research in education* (Vol. 24, pp. 355–392). American Educational Research Association.

Gordon, S., & Smith, E. (2021). Who are faculty assessment leaders? *Assessment & Evaluation in Higher Education*, 1–14.

Grainger, P. (2021). Enhancing assessment literacies through development of quality rubrics using a Triad based peer review process. *Journal of University Teaching & Learning Practice*, *18*(4), 4–14. doi:10.53761/1.18.4.4

Grainger, P., Purnell, K., & Zipf, R. (2008). Judging quality through substantive conversations between markers. *Assessment & Evaluation in Higher Education*, *33*(2), 133–142. doi:10.1080/02602930601125681

Grainger, P., & Weir, K. (2020). Creating quality rubrics through conversation. In P. Grainger & K. Weir (Eds.), *Facilitating student learning and engagement in higher education through assessment rubrics*. Cambridge Scholars Publishing.

Halliday, M. A. (2013). Meaning as choice. In L. Fontaine, Tom Bartlett, & G. O'Grady (Eds.), Systemic functional linguistics: Exploring choice (pp. 15-36). Cambridge University Press. doi:10.1017/CBO9781139583077.003

Halliday, M. A. K. (1993). Towards a language-based theory of learning. *Linguistics and Education*, *5*(2), 93–116. doi:10.1016/0898-5898(93)90026-7

Halliday, M. A. K., & Hasan, R. (1989). *Language, context, and text: Aspects of language in a social-semiotic perspective*. Oxford University Press.

Halliday, M. A. K., & Kress, G. R. (1976). *System and function in language: Selected papers*. Oxford University Press.

Halliday, M. A. K., & Matthiessen, C. (2014). *An introduction to functional grammar*. Edward Arnold. doi:10.4324/9780203783771

Hawe, E., Dixon, H., Murray, J., & Chandler, S. (2021). Using rubrics and exemplars to develop students' evaluative and productive knowledge and skill. *Journal of Further and Higher Education*, *45*(8), 1033–1047. doi:10.1080/0309877X.2020.1851358

Holmes, C., & Oakleaf, M. (2013). The official (and unofficial) rules for norming rubrics successfully. *Journal of Academic Librarianship*, *39*(6), 599–602. doi:10.1016/j.acalib.2013.09.001

Huang, S.-C. (2012). Like a bell responding to a striker: Instruction contingent on assessment. *English Teaching*, *11*(4), 99–119.

Hyland, K. (1996). Writing without conviction? Hedging in science research articles. *Applied Linguistics*, *17*(4), 433–454. doi:10.1093/applin/17.4.433

Jönsson, A., & Panadero, E. (2017). The use and design of rubrics to support assessment for learning. In D. Carless (Ed.), *Scaling up assessment for learning in higher education* (pp. 99–111). Springer. doi:10.1007/978-981-10-3045-1_7

Jönsson, A., & Svingby, G. (2007). The use of scoring rubrics: Reliability, validity and educational consequences. *Educational Research Review*, *2*(2), 130–144. doi:10.1016/j.edurev.2007.05.002

Knoch, U. (2009). Diagnostic assessment of writing: A comparison of two rating scales. *Language Testing*, *26*(2), 275–304. doi:10.1177/0265532208101008

Koh, K. H. (2011). Improving teachers' assessment literacy through professional development. *Teaching Education*, *22*(3), 255–276. doi:10.1080/10476210.2011.593164

Kohn, A. (2006). The trouble with rubrics. *English Journal*, *95*(4), 12–15. doi:10.2307/30047080

Lea, M. R., & Street, B. V. (1998). Student writing in higher education: An academic literacies approach. *Studies in Higher Education*, *23*(2), 157–172. doi:10.1080/03075079812331380364

Li, J., & Lindsey, P. (2015). Understanding variations between student and teacher application of rubrics. *Assessing Writing*, *26*, 67–79. doi:10.1016/j.asw.2015.07.003

Libarkin, J. (2008). *Concept inventories in higher education science*. Manuscript prepared for the National Research Council Promising Practices in Undergraduate STEM Education Workshop 2, Washington, D.C.

Lindström, L. (2006). Creativity: What is it? Can you assess it? Can it be taught? *International Journal of Art & Design Education*, *25*(1), 53–66. doi:10.1111/j.1476-8070.2006.00468.x

Macken-Horarik, M. (2008). *A "good enough" grammatics: Developing an effective metalanguage for school English in an era of multiliteracies. The ISFC*, Sydney.

Martens, K. S. (2018). How program evaluators use and learn to use rubrics to make evaluative reasoning explicit. *Evaluation and Program Planning*, *69*, 25–32. doi:10.1016/j.evalprogplan.2018.03.006 PMID:29660495

Martin, J. R., & Rose, D. (2008). *Genre relations: Mapping culture*. Equinox.

McDuff, N., Hughes, A., Tatam, J., Morrow, E., & Ross, F. (2020). Improving equality of opportunity in Higher Education through the adoption of an inclusive curriculum framework. *Widening Participation and Lifelong Learning : the Journal of the Institute for Access Studies and the European Access Network*, *22*(2), 83–121. doi:10.5456/WPLL.22.2.83

Medland, E. (2019). 'I'm an assessment illiterate': Towards a shared discourse of assessment literacy for external examiners. *Assessment & Evaluation in Higher Education*, *44*(4), 565–580. doi:10.1080/0 2602938.2018.1523363

Moni, R. W., Beswick, E., & Moni, K. B. (2005). Using student feedback to construct an assessment rubric for a concept map in physiology. *Advances in Physiology Education*, *29*(4), 197–203. doi:10.1152/advan.00066.2004 PMID:16298956

Montgomery, K. (2000). Classroom rubrics: Systematizing what teachers do naturally. *The Clearing House: A Journal of Educational Strategies, Issues and Ideas*, *73*(6), 324–328. doi:10.1080/00098650009599436

O'Donnell, J. A., Oakley, M., Haney, S., O'Neill, P. N., & Taylor, D. (2011). Rubrics 101: A primer for rubric development in dental education. *Journal of Dental Education*, *75*(9), 1163–1175. doi:10.1002/j.0022-0337.2011.75.9.tb05160.x PMID:21890846

O'Donovan, B., Price, M., & Rust, C. (2008). Developing student understanding of assessment standards: A nested hierarchy of approaches. *Teaching in Higher Education*, *13*(2), 205–217. doi:10.1080/13562510801923344

Oakleaf, M. (2009). *Writing rubrics right: Avoiding common mistakes in rubric assessment.* Association of College and Research Libraries 14th National Conference, Seattle, WA. http://meganoakleaf.info/writingrubricsright.pdf

Panadero, E., & Jönsson, A. (2013). The use of scoring rubrics for formative assessment purposes revisited: A review. *Educational Research Review*, *9*, 129–144. doi:10.1016/j.edurev.2013.01.002

Panadero, E., & Jönsson, A. (2020). A critical review of the arguments against the use of rubrics. *Educational Research Review*, *30*, 100329. doi:10.1016/j.edurev.2020.100329

Perkins, D. N., Jay, E., & Tishman, S. (1993). Beyond abilities: A dispositional theory of thinking. *Merrill-Palmer Quarterly*, *39*(1), 1–21.

Popham, W. J. (1997). What's Wrong—and What's Right—with Rubrics. *Educational Leadership*, *55*(2), 72–75.

Popham, W. J. (2009). Assessment literacy for teachers: Faddish or fundamental? *Theory into Practice*, *48*(1), 4–11. doi:10.1080/00405840802577536

Reddy, Y. M., & Andrade, H. (2010). A review of rubric use in higher education. *Assessment & Evaluation in Higher Education*, *35*(4), 435–448. doi:10.1080/02602930902862859

Richards, K., & Pilcher, N. (2014). Contextualising higher education assessment task words with an 'anti-glossary' approach. *International Journal of Qualitative Studies in Education : QSE*, *27*(5), 604–625. doi:10.1080/09518398.2013.805443

Sadler, D. R. (1998). Formative assessment: Revisiting the territory. *Assessment in Education: Principles, Policy & Practice*, *5*(1), 77–84. doi:10.1080/0969595980050104

Sadler, D. R. (2005). Interpretations of criteria-based assessment and grading in higher education. *Assessment & Evaluation in Higher Education*, *30*(2), 175–194. doi:10.1080/0260293042000264262

Sadler, D. R. (2011). Academic freedom, achievement standards and professional identity. *Quality in Higher Education*, *17*(1), 85–100. doi:10.1080/13538322.2011.554639

Sadler, D. R. (2013). Assuring academic achievement standards: From moderation to calibration. *Assessment in Education: Principles, Policy & Practice*, *20*(1), 5–19. doi:10.1080/0969594X.2012.714742

Sadler, D. R. (2014). The futility of attempting to codify academic achievement standards. *Higher Education*, *67*(3), 273–288. doi:10.100710734-013-9649-1

Schleppegrell, M. J. (2004). *The language of schooling: A functional linguistics perspective*. Routledge. doi:10.4324/9781410610317

Smith, C. D., Worsfold, K., Davies, L., Fisher, R., & McPhail, R. (2013). Assessment literacy and student learning: The case for explicitly developing students 'assessment literacy'. *Assessment & Evaluation in Higher Education*, *38*(1), 44–60. doi:10.1080/02602938.2011.598636

Stevens, D. D., & Levi, A. J. (2013). *Introduction to rubrics: An assessment tool to save grading time, convey effective feedback, and promote student learning*. Stylus Publishing, LLC.

Tan, K. H. K. (2020). *Assessment Rubrics Decoded*. Routledge. doi:10.4324/9780429022081

Tierney, R., & Simon, M. (2004). What's still wrong with rubrics: Focusing on the consistency of performance criteria across scale levels. *Practical Assessment, Research & Evaluation*, *9*(2), 1–7.

Torrance, H. (2007). Assessment as learning? How the use of explicit learning objectives, assessment criteria and feedback in post-secondary education and training can come to dominate learning. *Assessment in Education: Principles, Policy & Practice*, *14*(3), 281–294. doi:10.1080/09695940701591867

Troyan, F. J., Sembiante, S. F., & King, N. (2019). A case for a functional linguistic knowledge base in world language teacher education. *Foreign Language Annals*, *52*(3), 644–669. doi:10.1111/flan.12410

Turley, E. D., & Gallagher, C. W. (2008). On the" uses" of rubrics: Reframing the great rubric debate. *English Journal*, *97*(4), 87–92.

Walker, S., & Hobson, J. (2014). Interventions in teaching first-year law: Feeding forward to improve learning outcomes. *Assessment & Evaluation in Higher Education*, *39*(3), 326–338. doi:10.1080/0260 2938.2013.832728

Wang, W. (2017). Using rubrics in student self-assessment: Student perceptions in the English as a foreign language writing context. *Assessment & Evaluation in Higher Education*, *42*(8), 1280–1292. doi:10.1080/02602938.2016.1261993

Wiggins, G. (1998). *Educative sssessment. Designing assessments to inform and improve student performance*. Jossey-Bass.

Willis, J., Adie, L., & Klenowski, V. (2013). Conceptualising teachers' assessment literacies in an era of curriculum and assessment reform. *Australian Educational Researcher*, *40*(2), 241–256. doi:10.100713384-013-0089-9

Wingate, U. (2015). *Academic literacy and student diversity*. Multilingual Matters. doi:10.21832/9781783093496

KEY TERMS AND DEFINITIONS

Academic Literacy: An understanding of the way knowledge is created and communicated in a discourse community, the nature of relationships and interactions between participants in the discourse community, and the norms that regulate these interactions.

Analytic Rubric: An assessment tool to assess performance or learning on a given task. The analytic rubrics outlines criteria against which performance is assessed and various levels of quality at which performance may be demonstrated.

Assessment Literacy: An understanding of the purpose of assessment, how it fits into the learning trajectory, knowledge of the process of assessment, and the ability to evaluate work and identify means of improvement.

Criteria: Characteristics that are useful for determining the quality of work.

Discourse Community: A group of people who share the same values, goals, and language-use practices.

Performance Descriptors: Expressions of what a performance at a particular level looks like.

Rubric: An assessment tool that lists criteria for student work and articulates the levels of quality for each criterion.

Standards: Levels of academic achievement that students are expected to meet.

Teacher Development: The construction of teaching competences and pedagogical growth in teachers who make decisions in instruction and assessment.

Chapter 12
Self-Assessment:
Preservice Teachers' Concepts, Instruments, and Practices

Elsa Maria Ferro Ribeiro-Silva

Faculty of Sports Science and Physical Education, University of Coimbra, Portugal

Catarina Amorim

Faculty of Sports Science and Physical Education, University of Coimbra, Portugal

ABSTRACT

This chapter focuses on self-assessment as a decisive assessment for students' learning and offers the results of a study done with 72 university students. The authors investigated preservice teachers' views on self-assessment, the instruments used, and the implications for their students' learning. The results showed that while preservice teachers appear to understand what self-assessment is theoretically, in practice it is a non-systematic assessment with a planned day and criteria that is invariably quantitative. It appears that it is done because it is legally required and not because teachers believe students need to critically reflect on their learning.

INTRODUCTION

At a time when the teaching profession's credibility, identity, and professionalism are under constant attack, it is becoming increasingly necessary to professionally equip teachers with the theoretical, technical, practical, and ethical tools that will allow them to enter the teaching profession without turmoil.

In response to the Bologna Treaty (which created a European Space for Higher Education), to which Portugal was a signatory, initial teaching education was governed by Decree-Law No. 79/2014 (with reference to Decree-Law No. 74/2006 and Decree-Law No. 43/2007), which defines that the Teaching Practicum is a curricular unit of the last two semesters of the master's degree for teaching, providing preservice teachers with the experience of performing the teaching profession during an academic year.

Higher education institutions establish protocols with a network of cooperating schools (public or private) to host groups of two to four preservice teachers (the teaching practicum group) guided by a

teacher from the respective school (the school tutor), in collaboration with the faculty tutor, for the operationalisation of this curricular unit.

In addition to a procedural and formative evaluation of the preservice teachers, two formal moments of evaluation are established during the year of teacher training. The first aims to have all those involved (preservice teacher, school tutor, and university tutor) critically reflect on the preservice teachers' learning and on the entire respective learning process, with the goal of identifying and reinforcing their strengths and detecting difficulties to define strategies to overcome them. The second moment, which coincides with the end of the Teaching Practicum, has two goals: to identify skills that preservice teachers have not yet fully developed, providing them with clues to direct their future self-training and to contribute to a final grade. This classification is required for teachers to apply to the annual school placement throughout their careers.

The intermediate assessment is critical between these two moments because it allows us to intervene in the training process of a preservice teacher, adjusting or improving the corresponding training path while focusing entirely on the abilities already obtained and those that need to be taught in a different method, with no place for any type of classification.

BACKGROUND

Assessment, regardless of the form it takes, is one of the most challenging tasks for the novice teacher, as preservice teachers generally confirm in their final reports of the Teaching Practicum. Morgan and Hansen (2007) also identified this difficulty, stating that this is the area competence in which teachers feel less competent in their teaching practice.

Even though these issues persist throughout the school year, they are more focused at the start and are usually explained by the difficulty in articulating what they genuinely want to analyse and how to do it. Additionally, the problems of observation, the unfamiliarity of the pupils' identities, and, most importantly, the frequent use of unnecessarily complex and extensive assessment grids do not facilitate agile and quality data collection.

The preservice teachers' inexperience frequently causes them to apply endless grids, which they created or adopted from the disciplinary group, in which each activity is divided into several technical determinants with varying valuations. Because the sum of parts, that is, the sum of each rated technical gesture, does not direct us to the true worth of the pupils' performance, this frequently leads to an evaluation that is completely different from reality. Due to its decontextualisation and withdrawal of decision-making capacity, this assessment by isolated technical gestures does not indicate the pupils' performance level in the sport in question.

These obstacles are summarised in the final report by one of this year's preservice teachers: *the primary issues in the formative assessment process in the early stages of the teaching practicum were in the coherence of gathering information on the performance of our students, as our observation capacity did not correspond to the quantity of criteria to be observed.* Adding, subsequently, that *the instruments used in the assessment were designed and authorised by the school's physical education group and that the use of unfamiliar tools made this procedure challenging*.

On the other hand, becoming a teacher at a school where most of the other teachers still see assessment and classification as synonymous concepts and use a traditional assessment system enables preservice teachers to easily fall into the same cycle. In summary, one cannot apply something that one has never

experienced as a student, has never been instructed in, and has never seen anyone perform (Lorente & Kirk, 2013; Pérez-Pueyo et al., 2017).

Despite the training in pedagogical assessment offered to preservice teachers in the first year of the Master's, their lack of consolidation combined with the influence of the teachers at the school where they are placed almost always leads to a traditional assessment focused on classification, regardless of the type and moment of assessment they intend to carry out, although they advocate formative assessment in theory.

Calatayud stated in 2007 that it was critical to deepen future teachers' training in pedagogical assessment so that they stopped thinking of it as something specific to the teaching process. The author claims that to certify, through classification, the level of learning achieved by students, making it coincide with concepts such as measurement, classification, and qualification, these functions do not represent the primary actions for which assessment should serve.

Much later, the International Association for Physical Education in Higher Education (2021) took a position on assessment in Physical Education, arguing that in initial teacher education, a high-quality education for future PE teachers must be ensured for them to develop assessment literacy and learn to design and implement a reliable, valid, authentic, and transparent assessment that involves students.

This position, while important, is not novel, as several authors such as Boud (2010) and Lorente-Catalán (2015) have previously argued that assessment is more effective when students and tutors/teachers share responsibility for it, implying the use of participatory practices such as self-assessment, peer assessment, or co-assessment, which must be incorporated into the design and planning of units of work if we are to fully engage in their own learning process in order to promote the kind of skills that they need to become lifelong learners. As a result, whenever the student is involved in this process, that is, whenever they self-assess, assessment plays a fundamental role as a regulator of the educational process and as a privileged moment of learning, with self-assessment being one of the richest forms of assessment (Black & William, 2003; Calatayud, 2008; Lopéz-Pastor, 1995).

According to Calatayud (2008), that is the best strategy for teaching responsibility and autonomy, as well as learning to value, criticise, and reflect on the teaching and learning process, as a learning instrument that serves both those who learn and those who teach, in the sense that it helps their students learn better.

Given that self-assessment entails becoming aware of what one is doing and the goals one wants to achieve in order to improve it, Calatayud (2008) defends four critical principles on which self-assessment should be based, which include recognising: i) the need; ii) how it differs from self-classification; iii) that students are capable of performing it; and iv) the importance of the teacher considering the pupils' self-assessment when classifying each student.

That is, when teachers suggest self-assessment to their pupils, they must aim to make them aware of their learning progress, assist them in taking responsibility for the development of their activities through the development of autonomy and self-management, and be a motivator and reinforcer of learning.

Brown and Harris (2013) also reinforced this idea, claiming that self-assessment is a process of self-perceptions of oneself as a person and thus of one's views, values, feelings, reactions, and qualities. Learning from the self-assessment process extends beyond the classroom's space and time to become lifelong learning in the context of personal and professional performance. The self-assessment process allows for the use of metacognitive processes, allowing for the construction of learning (Moreno et al., 2013). As a result, self-assessment entails a set of decisions made by the individual and his responsibility for them, as well as the positive and negative repercussions that ensue, showing the need for including

it in the training process of individuals, facing it like any other teaching-learning material (Brown & Harris, 2013; Yan, 2019).

Through this process, which Zimmermann et al. (2012) characterise as self-regulation, the student can acquire autonomous and proactive learning in which they can construct assertive methods based on their personal qualities, emotions, motives, and cognitions. They will be able to develop assertive strategies that, in turn, will maximise their learning.

According to Flores (2020), initial teacher education is a place where new teachers build, deconstruct, and reconstruct their perspective of being a teacher, and the Teaching Practicum is considered the primary learning moment in the teaching profession (Cardoso et al., 2016; Flores, 2020; González-Calvo et al., 2020), as it is the first contact with the reality of being a teacher and the first opportunity that students have to develop their practice in a real teaching setting.

MAIN FOCUS OF THE CHAPTER

This chapter describes a study whose objective was to identify preservice teachers' perceptions and execution of self-assessment (concepts, instruments, and practices). The participants were part of a master's degree in Physical Education Teaching at a public university in Portugal, in the intermediary phase of the school placement, to intervene and improve the participants' knowledge during the second part of the academic year.

Issues, Controversies, Problems

As previously stated, preservice teachers are responsible for teaching a third cycle or secondary school class for an entire academic year, as well as a variety of non-teaching responsibilities intrinsic to the role of teacher, acquiring the title of 'preservice teacher' at the school where they teach.

As the preservice teachers each take a class from the school tutor schedule, the last must attend all classes because they are the holder of the class and the actual teacher. This encourages close mentoring and even the creation of an affective relationship between preservice teachers and school tutors, with many benefits but some drawbacks, such as continual consultation by preservice teachers rather than seeking out current bibliographic sources or other types of self-training. The scarce research tradition of the teaching staff is negatively contributing to teachers being agents of innovation and change (Calatayud, 2018), leading the preservice teachers, in their inexperience, to be dragged into this cycle.

This reality causes preservice teachers to uncritically replicate their school tutors' ideas, forcing them to enter a traditional teaching methodology that is very comfortable in our schools rather than being the source of much-needed pedagogical innovation and renewal. This is the only way to explain why, when we asked a preservice teacher if they perform self-assessments of their students, the answer was invariably affirmative; however, when we asked what their objective was with that assessment and how they mobilised the results obtained, the answers were not as assertive and clear.

This situation prompted us to investigate preservice teachers' views of self-assessment, the instruments they use to conduct such assessment, and the implications for the learning and classification of the respective students. To do this, we decided to ask the 72 preservice teachers from the 2021-2022 academic year (21 female and 49 male, aged between 22 and 43 years old) a series of eight open-ended

questions on the self-assessment process about what they had done (or not done). We were guided by the evaluation questions (What? When? Why? How? and For what?) (Table 1).

Table 1. Open-ended questionnaire

Assessment questions	Question	Desired information
What?	1. What is self-assessment for you?	Understand the existing self-assessment concept(s)
When?	2. Did you self-assess your students? When (time(s) of the school year)?	Understand when self-assessment is applied
Why?	3. What purpose did you have when you made it?	Understand what was intended with the self-assessment
How?	4. How did you do it? Describe the form (oral, written, etc.), instruments (open-ended questionnaire, closed-response questionnaire, etc.), location (gym, classroom, etc.). 5. Did you analyse the data obtained (students' responses)?	Understand the implementation process
For what?	6. What did you do with that data? 7. What advantages do you think self-assessment has for students who do it? 8. What advantages do you think carrying out this self-assessment process had for you as a preservice teacher?	Perceive the recognition that is given to that form of assessment (for students and for preservice teachers)

Source: the authors

Those questions were placed by chance at the end of the respective intermediate self-assessment grid of their own performance as preservice teachers. This grid was submitted on the Teaching Practicum curricular unit's digital platform during a period designated for intermediate self-assessment, around the middle of the school year.

An open-ended questionnaire was chosen to 'give a voice' to respondents, covering the following self-assessment characteristics: i) concept of self-assessment; ii) typology of instruments (oral questioning; written questionnaire with close response; written questionnaire with open response); iii) focus of instruments (incidence of evaluation); iv) potential of the instruments (value or interest of the data obtained). We conducted this research using a qualitative approach under the constructivism paradigm (Creswell & Creswell, 2018).

A semantic analysis of the content of the responses was carried out (Bardin, 2011), assuming the questions as starting categories. This analysis was done in parallel by two researchers who maintained a continual debate among themselves, achieving agreements based on discussions on the interpretation of the preservice teachers' discourses and their associated codification and classification (Creswell & Creswell, 2018), to protect the research's credibility. When comparisons between them failed to reveal new linkages and characteristics, it was thought that data saturation had occurred. In this manner, the data's representational richness was judged exhausted in a cyclical analytical process (Flick, 2007). To summarise the findings of the content analysis performed on the results, we have:

1.   What is self-assessment for you?

With this question, we aimed to comprehend the notions or concepts of preservice teachers' self-assessment, which appeared ambiguous and were centred on four unique individualities: 1) student; 2) teacher; 3) the preservice teacher; and 4) classification.

Depending on the student, they may be given the central role in a reflection process, although this is not always the case. Even though a significant majority refers to the reflection on the students' performance in the classes regarding the contents of the unit of work, there were still several references to the reflection on the respective performance, behaviour, or even the learning that was 'transmitted'. Many of the preservice teachers believe it is important to base self-assessment on a numerical scale.

Self-assessment refers to a student's ability to appraise their own performance, whether in terms of motor level, commitment, or knowledge of the Physical Education subject. (PST 13)

Self-assessment is a time when students reflect on their classes and their attitudes and actions in class, allowing them to rate each item (behaviour, technical actions, punctuality, etc.). We propose to analyse students' performance in classes in the many domains of the subject through self-assessment. (PST 22)

Despite frequent references to self-assessment as 'the moment of reflection on their strengths and weaknesses', there is no clear association or understanding of that reflection as a decisive and integral factor of the learning process, placing the student in a situation of reflection on the learning process over which they have no decision-making power.

It is a way of ensuring that students reflect on their process and on all the learning that was transmitted to them, understanding what was or was not retained by the students. (PST 25)

This idea is reinforced when nearly a quarter of the preservice teachers associated their students' self-assessment not only with classification but also, and most importantly, with the confirmation of the grade they had reserved for them, while insisting that that form of assessment represents a reflective process of the students.

Self-assessment is a reflective time for students in which they independently reflect on what they have accomplished and their dedication and qualify these attitudes with a numerical value that may or may not match the teachers' opinion. (PST 26)

When self-assessment is successfully implemented and students understand what they are being assessed on, the process becomes easier since the students' impression of evaluation becomes a mechanism for the instructor to compare what their assessment is. (PST 34)

Self-assessment is a tool that allows students to reflect on their behaviour in physical education sessions. This is, without a doubt, a sort of self-reflection, with the goal of allowing the student to locate themselves in themself, identifying their strengths while developing measures to reduce their flaws. (PST 17)

Centred on the teacher, it is a pedagogical tool for getting to know the pupils and guiding them in their pedagogical decisions.

Self-assessment is a useful tool that allows students to reflect on their performance while meeting previously established goals. It is a tool that allows the teacher to not only comprehend the student's opinion of himself but also to make the student reflect on their actions in Physical Education lessons. (PST 17)

Self-assessment serves the purpose of allowing the teacher to understand the student's perception of their own performance as well as the student's capacity for self-reflection. (PST 37)

Centred on the preservice teacher, defining self-assessment as an action taken by themselves on the work they had developed, and that action is often associated with expressions such as self-reflection, regulation, judgment, assessment, self-criticism, retrospection, critical opinion, or critical spirit as synonyms, revealing the divergence in their perceptions of self-assessment.

(…) the assessment assumes a self-regulation function, specifically knowledge or learning. In this scenario, self-assessment is a process that allows for the establishment of a student's awareness of learning, in which students examine the balance of the work completed to reflect and analyse their performance throughout the process. (PST 12)

In my opinion, self-assessment is a judgement on our knowledge and/or behaviours. When this process is repeated on a regular basis, it allows for the measurement of evolution and can also serve as a booster for increased motivation and goal definition. (PST 16)

For me, self-assessment is an individual's ability to evaluate oneself constructively or adversely. Furthermore, I believe it is a way that seeks to help us consider what we need to do to improve our performance on tasks. (PST 27)

Centred on the grade, where, although continuing to characterise self-assessment as a reflection, the ultimate goal is to arrive at a value, a grade, and is thus a classifying action.

Self-evaluation is a period for self-analysis, awareness, and reflection on the student's performance across all assessment points. It is a moment in which it is placed quantitatively on a scale of 1 to 5 in relation to its performance. (PST 6)

I used co-participation assessment sheets in some subjects where students recorded: times and distances (track and field); movements performed correctly (gymnastics), according to the critical components of each technical gesture; number of movements performed (volleyball and tennis); and so on. The students completed the self-assessment at the end of the first period, with the grade they thought they deserved and a record of their primary problems and strengths, as well as strategies to improve their weaknesses. (PST 58)

2.   Did you self-assess your students? When (time(s) of the school year)?

The preservice teachers all agreed on the answers to both questions because they all claimed to have completed self-assessment at the end of a teaching cycle (units of work, term, semester, etc.). Many were

those who saved this assignment for the last lesson of the school year, even noting the precise day, not recognising that this is not the goal of self-assessment and hence not of significance.

Self-assessment was completed at the end of the first period, precisely in the last lesson on December 16, 2021. (PST 2)

The students' self-evaluation has been completed at the end of each teaching unit, or every two weeks. The self-assessment is also done at the end of the semester, when students reflect on their whole semester. (PST 9)

Yes, students are assessed three times: at the completion of each unit of work, at the end of each session, and at the end of the academic year. (PST 24)

On January 24, 2021, the last lesson of the first semester. (PST 63)

3.   What purpose did you have when you made it?

The answers did not reveal any intentionality in carrying out the self-assessment, with no defined purpose, with the majority noting that with that type of assessment they intended to inspire students to reflect on their performance in class, but differing in the parameters considered, with speeches focused on technical and tactical execution and others on attitudes and values.

The main goal of the first term self-assessment was to make students reflect on their dedication in class, on their action in practice, on the learning received from the technical and tactical execution of the exercises conducted in class, and on the knowledge acquired from the subjects taught. (PST 55)

To learn about the students' perceptions of their own performance as well as their motor and technical/ tactical skills in each subject's content. (PST 56)

Most of the time, reflection and, subsequently, self-assessment are used to compare the grade that the students give themselves with the grade that preservice teachers have already determined to give them, with no justification for this comparison.

I wanted to compare my assessment to my pupils' perceptions of their performance over the term. (PST 2)

The goal is to determine whether each student develops a critical and rational reflection on their behaviour and progress in teaching and learning circumstances or whether they simply give a grade just because, without justification. (PST 6)

The main objective of using self-assessment is to give students a chance to reflect on their performance during the unit of work in the hope that they will reflect critically on their actions. Furthermore, it is critical to determine whether the classification that pupils believe they deserve corresponds to the classification that the teacher believes they should receive. (PST 8)

I intended to detect differences or similarities between the teacher's and students' grades. Another important goal is for students to reflect on their performance and identify areas for improvement. (PST 9)

We found no evidence in any of the speeches that the student would be given a voice in the event of divergence or that that evaluation had any bearing on the teacher's final classification. Similarly, it is unclear how self-assessment results will be mobilised to improve students' learning processes. This reality reveals that, contrary to what Calatayud (2008) claims, the adoption of self-assessment by preservice teachers for their students occurs because of legal imposition or school standards.

4. How did you do it? Describe the form (oral, written, etc.), instruments (open-ended questionnaire, closed-response questionnaire, etc.), location (gym, classroom, etc.).

We wanted to understand not only the process of executing the self-assessment but also the way it was carried out and the instruments used; therefore, we asked this question.

I did self-assessment orally, asking students to indicate a value ranging from 0 (terrible) to 5 (outstanding) inside the two units of work, corresponding to the value they believe they deserve in terms of behaviour and commitment. This self-assessment was completed in the gym, where we generally hold the session (a place that makes students more comfortable answering honestly and makes them reflect on what they did there). (PST 26)

I accomplished it in two stages. The first was written, and the student only had to name and justify the grade they desired. The second was an oral moment in which, according to the final assessment grid, it inquired about the student's commitment and performance, prompting them to reflect on their attitudes and learning in each parameter, before clarifying the student's grade and justifying it with moments that occurred in class. (PST 55)

In the gym, I used written and oral forms, with closed-ended questionnaires, as well as a dialogue with students to see if their vision matched mine. (PST 65)

As for the type of instruments used for self-assessment, approximately one-third of the preservice teachers used a digital support instrument, which we believe was influenced by the COVID-19 pandemic as well as current concerns about the promotion of the use of technologies in school and in physical education classes.

Informal self-assessment is done orally and is frequently done in class, either during or after the final assessment. The formal self-assessment, which took place at the end of the first term, consisted of responding to an online questionnaire. (PST 18)

The students completed an existing questionnaire the school uses on its digital platform. (PST 46)

Regardless of the sort of assistance, more than half of the preservice teachers chose to use a questionnaire with closed questions, almost all of which were built by levels and for which students merely had to mark an option.

The self-assessment was created using an online form (Google Forms), which allowed students to complete it during class, anywhere, and at any time during the week of December 13 to 17. The form consisted of 29 closed questions. (PST 17)

Written, using closed responses in Google Forms according to the Likert scale (1, 2, 3, 4, 5). (PST 28)

The self-assessment was done in writing, using closed-response questionnaires, with a rating from 1 to 5 for each question. (PST 36)

5. Did you analyse the data obtained (students' responses)?

Only one of the 72 students revealed that they did not analyse the self-assessment data. However, when they describe the way they did it, we realised that they were characterised by being quantitative (assessment for classification) and, once again, were analysed considering the grade that the preservice teacher had assigned or was about to assign to the student, which the preservice teacher considered to be the correct grade.

Many preservice teachers also state unequivocally that their grade is unquestionable and that, if they disagree with the student's approach, they are required to explain why it is incorrect. In other words, the idea of reason being on the side of the self-assessed was never put forward.

I analysed it, yes. After each answer, I wrote down each student's grade after each answer, while also conducting a quick discussion of the grades and asking what they were based on for assigning that grade. (PST 39)

Yes, each self-assessment form is examined after the session, and I provide feedback on the same scale that the students filled out, as well as verbal feedback in the following class in cases where there was a significant difference between the students' perception and my perception. (PST 49)

Quantitatively, yes. (PST 51)

Yes, in comparison to the grades I assigned them. (PST 54)

6. What did you do with that data?

The responses to this question were summarised into three categories: i) those who did nothing with the data; ii) those who saved them to be compared with the findings of other self-assessment times (end of term or unit of work); and iii) those who promptly compared with their classification. In short, regardless of the outcome, the students' learning process was unaffected. As a result, while some preservice teachers claim that the results helped them adjust their pedagogical decisions, the vast majority carried out the self-assessment using a questionnaire with closed questions and a likert scale, which, contrary to their beliefs, cannot have been of great value in the learning process of their students.

In the case of the form filled out by the students during the first term, I admit that I did not analyse the data obtained and, as a result, did not appraise them. As I previously stated, the form is incredibly broad, if not difficult to interpret, and does not provide enough particular data for research. (PST 5)

The data were retained at the end of the school year to assess if the students' reflections reached or did not approach the teacher's opinion as the year proceeded and the student-teacher relationship developed. (PST 8)

To compare the data, I put it in a column next to the teacher's grade. (PST 9)

7.   What advantages do you think self-assessment has for students who do it?

Surprisingly, the responses to this question suggest that preservice teachers have a very good idea of the purpose of this type of evaluation, noting that it is a reflection by the student on their performance to enhance it. Many of them, however, continue to believe that it should be procedural, arguing that self-assessment is useful since it allows the results to be mobilised later in a following unit of work or in a following term.

It challenges students to think about what they have done, where they did well, and where they could improve in the next unit of work or term. (PST 2)

The fact that students actively participate in their teaching-learning process through cognitive processing in relation to their practical performance is a significant advantage. Students are prompted to ponder and recognise their strengths, possibilities for progress, and, later, the paths they need to take to overcome the challenges. (PST 5)

The idea was also popular that the reflection should not focus on the students' performance but on the evaluation factors.

It helps the student to deeply reflect on the entire set of evaluation parameters so that they do not focus solely on their physical capabilities but also on their behaviour, commitment, attitudes, and beliefs, because the sum of everything is what determines the final grade. (PST 1)

In my opinion, self-assessment is beneficial because it allows students to reflect on the teaching-learning process that they were involved in during that term or school year and then determine where they need to improve, what they can change to overcome certain physical aspects, and their attitudes or behaviours in class. (PST 33)

8.   What advantages do you think carrying out this self-assessment process had for you as a preservice teacher?

Finally, when we asked them about the benefits they saw in self-assessment for their training, not all of them focused on the value of that type of assessment for themselves as preservice teachers, with some emphasising its importance for their students.

As a preservice teacher, I believe that this process allowed me to obtain more knowledge and ideas about what students think, that is, what they define as the most important parameters during classes, which will help me in the future to define the main assessment parameters in a more assertive manner so that there are no significant differences during self-assessment. (PST 1)

It surely aids the teachers' comprehension of whether the information was consolidated by the students, where the doubts are and how we can resolve them, where the students believe something is missing to ensure their success, and how the teacher is perceived by each student. (PST 25)

This self-assessment provided me with a critical reflection on my own performance as a preservice teacher, recognising and identifying my strengths and weaknesses, looking for improvement strategies for these same weaker points, and allowing me to be aware of my performance level and my improvements since the beginning of the school year. (PST 69)

Unsurprisingly, there were numerous remarks that continued to emphasise the importance of self-assessment for comparison with the grades that the preservice teachers imagined they would give and, thus, for the classification of the various students, which was transversal to all questions.

This process allowed me to identify that students do not always appropriately perceive their level. In one specific situation, towards the end of the term, I found that some students believed their grade was much higher than it was, which led me to believe that they were still unable to acknowledge their challenges. In this regard, conducting the self-assessment allowed me to clarify the students' performance and reaffirm what is still required to achieve this level. (PST 30)

I was able to obtain the students' point of view through self-assessment, verifying if they can regulate their learning and having the perception that students are aware of their abilities. Many of them who have a good performance put grades on the self-assessment sheet that they do not correspond with, not knowing how to value their work, and the opposite also happens: students who have an unsatisfactory performance believe that they deserve a good grade. (PST 32)

SOLUTIONS AND RECOMMENDATIONS

As we started by saying, this study developed from a distrust about the unclear understanding and poorly consolidated training in pedagogical assessment in initial teacher education, which appears to be substantiated by the results. Despite Calatayud's claims in 2008, when the development of master's degrees in teaching was near and enormous hopes were placed on better training in pedagogical assessment in those new teaching cycles, the truth is that this is not very evident.

In contrast to Fernandes (2021), who argues that assessment must accompany all pedagogical practices and, in particular, teaching processes in order for them to be continuously and systematically regulated and

improved, the findings show a self-assessment understood by preservice teachers as having the function of confirmation, classification, or balance of a teaching cycle that has already ended while continuing to define it as a process of reflection.

This evaluation, carried out at the end of anything, is devoid of value because it prevents the results from being integrated into the learning process. According to Dochy et al. (2002), students should be considered active participants who share responsibility in the process of self-assessment, reflection, and collaboration in a continuous dialogue with the teacher. However, teachers and preservice teachers continue to regard assessment as a task that occurs after everything else has been completed and in which students have no say in something that directly concerns them. This viewpoint leads to teaching methods and, as a result, 'traditional-summative' assessment (Hernán et al., 2019).

Calatayud (2007) stated more than fifteen years ago that many factors determined the persistence of an evaluative culture as a sanctioning and classifying instrument, highlighting three of them: i) the legacy of an 'examiner' evaluation; ii) the need for a school culture that integrated the value and meaning of assessment as a critical learning activity and professional and institutional development; and iii) a lack of basic training in formative assessment.

The insufficient length of the curricular units dedicated to evaluation training, followed by an academic year of Teaching Practicum immersed in physical education teachers' groups where that training is not normally reflected or updated, means that future teachers will be unable to abandon the vision of self-assessment for classification and, thus, to be carried out at the end of any teaching cycle.

Lorente-Catalán and Kirk (2016) concluded in their study with physical education preservice teachers, such as this one, that while preservice teachers revealed a willingness to use new forms of assessment in their future professional practice, which puts it at the service of student learning, there are several recent studies in the field of assessment in physical education (Hay & Penney, 2013; Lopez-Pastor et al., 2013) that highlight the barriers and resistance that novice teachers face. Almost all because of a lack of pedagogical assessment research and teachers' willingness to leave the safety and comfort of what they already know how to do, even if the outcomes are not remarkable, for something novel, demanding, and promising for themselves, preservice teachers, and students' education.

The findings indicated diverse and ambiguous notions of self-assessment among preservice teachers. Self-assessment moments were nearly always placed at the end of each academic term, when pupils were classified, typically summed up in a numeric value provided by the student as a categorisation they considered they merited, but with no bearing on the grade previously designed for them by the preservice teacher.

Furthermore, the instruments used were frequently small questionnaires with closed questions that only asked the student to put the classification they thought they deserved, or the same question but put orally by the preservice teachers and whose value, advanced by each of the students, the preservice teachers oversaw registering. Few instruments were mobilised that demanded self-reflection on one's own performance.

These findings appear to indicate a lack of recognition of the role and value of self-assessment on the part of the preservice teachers, which has gone unreported by the respective schools and faculty tutors, possibly because they too do not value self-assessment in initial teacher education.

Knowing that in students' conceptions of assessment and self-assessment in classes, built from their experiences, students associate assessment with the end of a process, striving for the highest ratings at the expense of better learning (Vieira, 2013), this way of acting by the preservice teacher does not

contribute to the deconstruction of those conceptions. According to Vieira (2013), it is the students' obligation to be involved in their own learning.

This is stated unequivocally in the Profile of Students Leaving Compulsory Schooling (approved by Order No. 6478/2917 of the Portuguese Republic), which states that students should be able to recognise their strengths and weaknesses, express their needs, and seek the most effective help and support to achieve their goals at the end of the 12 years of compulsory education. They must also be able to plan, implement, and evaluate methods on their own to meet the goals and tasks they set for themselves.

FUTURE RESEARCH DIRECTIONS

Assessment in education, despite being studied and having a significant existing bibliographic production, continues to be an ever-present issue. The (re)emergence of active approaches, which advocate for increased student engagement and autonomy, needs reflection on the assessment process, specifically self-assessment, and its teaching to students of all levels of education. Based on this hypothesis, we believe that the natural next step for this study will be to replicate it for these pupils at the conclusion of the school year, allowing us to determine whether the techniques we outlined based on the findings here had favorable outcomes.

We also intend to continue with a comparative study of how trainees classified at the level of excellence and those classified at the level of sufficient see and implement self-assessment, with the goal of obtaining data on their learning process and skill development.

CONCLUSION

Recalling the objective of this study, it was intended to identify preservice teachers' perceptions and execution of self-assessment (concepts, instruments, and practice) in the intermediary phase of the school placement to intervene and improve the participants' knowledge during the second part of the academic year.

The results guide us towards a self-assessment performed at the end of the teaching and learning process and disconected from it. This assessment is almost always quantitative in nature and with no relation with the learning goals, which are unknown to the pupils.

It appears to us that self-assessment is carried out because it is legally required and not because teachers think that students need to critically reflect on their learning (both in terms of process and level attained) as a way of becoming aware of what they need to learn and being guided by that.

Self-assessment activities had a favourable and substantial influence on students' academic performance, according to Yan et al. (2021). But how can it be done with a non-systematic self-assessment, a deadline, no specific criteria, and nearly invariably a quantitative approach?

As a result, and in accordance with Black and William (2003), who state that self-assessment learning should not be overlooked in its improvement through a process that requires time, practice, and intentionality, we consider these results critical to guiding the teaching process in initial teacher education and, more specifically, in the Teaching Practicum.

If evaluation is to be an educational process, students must be trained to evaluate and improve their own performance through self-assessment (Siedentop & Tannehill, 2000).

Faced with an uncertain and continuously changing educational landscape, self-assessment competency is becoming increasingly important, making training in self-assessment processes critical for improving educational practices (Calatayud, 2018).

This was the motivation for doing this study in the middle of the school year, allowing us to work to modify the preservice teachers' views of self-assessment, but more importantly, of pedagogical assessment, while also anticipating that resistance to change will be significant.

REFERENCES

Black, P., & Wiliam, D. (2003). 'In Praise of Educational Research': Formative assessment. *British Educational Research Journal*, *29*(5), 623–637. doi:10.1080/0141192032000133721

Boud, D. (2010). *Seven Propositions for Assessment Reform in Perspective*. Routledge.

Brown, G. T. L., & Harris, L. R. (2013). Student self-assessment. In J. H. McMillan (Ed.), *The SAGE handbook of research on classroom assessment* (pp. 367–393). doi:10.4135/9781452218649.n21

Calatayud, M. A. (2007). La evaluación como instrumento de aprendizaje y mejora. Una luz al fondo. In A. M. Calatayud (Coord). La evaluación como instrumento de aprendizaje. Estrategias y técnicas. (p. 9 – 23). Madrid.

Calatayud, M. A. (2008). Establecer la Cultura de la Autoevaluación. *Padres y Maestros*, *314*(febrero), 30–34.

Calatayud, M. A. (2018). La autoevaluación. Una propuesta formativa e inovadora. *Revista Ibero-Americana de Educação*, *72*(2), 135–152. doi:10.35362/rie7623081

Creswell, J., & Creswell, J. D. (2018). *Research Design: Qualitative, Quantitative, and Mixed Methods Approaches* (5th ed.). Sage Publications.

Decreto-lei n.º 43/2007 de 22 de fevereiro. Diário da República n.º 38/2007 - Série I de 2007-02-22, 1320-1328. Lisboa: Ministério da Educação. https://data.dre.pt/eli/dec-lei/43/2007/02/22/p/dre/pt/html

Decreto-lei n.º 74/2006 de 24 de março. Diário da República n.º 60/2006 - Série I-A de 2006-03-24, 2242-2257. Lisboa: Ministério da Ciência, Tecnologia e Ensino Superior. https://data.dre.pt/eli/dec-lei/74/2006/03/24/p/dre/pt/html

Decreto-lei n.º 79/2014 de 14 de maio. Diário da República n.º 92/2014 - Série I de 2014-05-14, 2819-2828. Lisboa: Ministério da Educação e Ciência. https://data.dre.pt/eli/dec-lei/79/2014/05/14/p/dre/pt/html

Dochy, F., Segers, M., & Dierick, S. (2002). Nuevas Vías de Aprendizaje y Enseñanza y sus Consequencias: una Nueva Era de Evaluación. *Revista de Docencia Universitária, 2*(2) Education: Classroom Teachers' Perspective. *The Journal of Educational Research*, *101*(2), 99–108. doi:10.3200/JOER.101.2.99-112

Fernandes, D. (2021). *Para uma Fundamentação e Melhoria das Práticas de Avaliação Pedagógica. Projeto de Monitorização Acompanhamento e Investigação em Avaliação pedagógica*. Universidade de Lisboa - Instituto de Educação, 3.

Flick, U. (2007). Managing quality in qualitative research. *Sage (Atlanta, Ga.)*.

Flores, M. A. (2020). Feeling like a student but thinking like a teacher: A study of the development of professional identity in initial teacher education. *Journal of Education for Teaching*, *46*(2), 145–158. doi:10.1080/02607476.2020.1724659

González-Calvo, G., Varea, V., & Martínez-Álvarez, L. (2020). 'I feel, therefore I am': Unpacking pre-service physical education teachers' emotions. *Sport Education and Society*, *25*(5), 543–555. doi:10.1080/13573322.2019.1620202

Hay, P., & Penney, D. (2013). *Assessment in Physical Education: A Socio-cultural Higher Education*. Australian Learning and Teaching Council., Available at http://www.uts.edu.au/sites/default/files/ Assessment-2020_propositions_final.pdf

Hernán, E., López-Pastor, V., & Pérez-Brunicardi, D. (2019). Por qué hago evaluación formativa y compartida y/o evaluación para el aprendizaje en EF? La influencia de la formación inicial y permanente del profesorado. *Revista RETOS*, *36*, 37–43.

International Association for Physical Education in Higher Education (2021). Tomada de Posição Sobre Avaliação em Educação Física. *Boletim SPEF, 42*.

López-Pastor, V., Kirk, D., Lorente-Catalán, E., MacPhail, A., & Macdonald, D. (2013). Alternative assessment in physical education: A review of international literature. *Sport Education and Society*, *18*(1), 57–76. doi:10.1080/13573322.2012.713860

Lorente-Catalán, E., & Kirk, D. (2015). Student teachers' understanding and application of assessment for learning during physical education teacher education course. *European Physical Education Review*, *22*(1), 65–81. doi:10.1177/1356336X15590352

Moreno, A., Trigueros, C., & Rivera, E. (2013). Autoevaluación y Emociones en la Formación Inicial de Profesores de Educación Física. *Estudios Pedagógicos (Valdivia)*, *39*(1), 165–177. doi:10.4067/S0718-07052013000100010

Morgan, P., & Hansen, V. (2007). Recommendations to Improve Primary School Physical Education: Classroom Teachers' Perspective. *The Journal of Educational Research*, *101*(2), 99–108. doi:10.3200/JOER.101.2.99-112

Vieira, I. (2013). *A autoavaliação como instrumento de regulação da aprendizagem*. [Dissertação apresentada à Universidade Aberta]

Yan, Z. (2019). Self-assessment in the process of self-regulated learning and its relationship with academic achievement. *Assessment & Evaluation in Higher Education*, *45*(2), 224–238. doi:10.1080/02602938.2019.1629390

Yan, Z., Wang, X., Boud, D., & Lao, H. (2021). The effect of self-assessment on academic performance and the role of explicitness: A meta-analysis. *Assessment & Evaluation in Higher Education*. Advance online publication. doi:10.1080/02602938.2021.2012644

Zimmermann, P., Flavier, É., & Méard, J. (2012). L'identité professionnelle des enseignants en formation initiale. *Spiral-E. Revue de recherches en éducationι, 49*, 35-50. https://doi.org/ doi:10.3406/spira.2012.1724

ADDITIONAL READING

Allal, L. (2020). Assessment and the co-regulation of learning in the classroom. *Assessment in Education: Principles, Policy & Practice*, *27*(4), 332–349. doi:10.1080/0969594X.2019.1609411

Cañadas, L. (2022). Procesos de auto-evaluación y co-evaluación en educación física. Una revision sistemática. *Revista Iberoamericana de Evaluación Educativa*, *15*(1), 161–176. doi:10.15366/riee2022.15.1.009

Darling-Hammond, L., Burns, D., Campbell, C., Lin Goodwin, A., Hammerness, K., Low, E., McIntyre, A., Sato, M., & Zeichner, K. (2017). *Empowered Educators: How High-Performing Systems Shape Teaching Quality Around the World*. Jossey-Bass.

Fletcher, A. (2018). Assessment to develop students' strategies and competence as learners. In M. Barnes, M. Gindidis, & S. Phillipson (Eds.), *Evidence-Based Learning And Teaching* (pp. 123–137). Routledge. doi:10.4324/9781351129367-11

Gariglio, J. (2021). Beginning Physical Education teachers' induction: Discoveries and survival in the profession. *Sport Education and Society*, *26*(7), 733–745. doi:10.1080/13573322.2021.1884061

Liu, J., Xiang, P., McBride, R., & Chen, H. (2020). Self-regulated learning strategies and achievement goals among preservice physical education teachers. *European Physical Education Review*, *26*(2), 375–391. doi:10.1177/1356336X19859602

López-Pastor, V. M., Fuentes-Nieto, T., & Jiménez-Herranz, B. (2020). Evaluación formativa, compartida y auténtica en educación física. *Tándem. Didáctica de la Educación Física*, *69*, 7–14.

Moura, A. (2022). *From assessment for ranking toward assessment for learning: An action research study in preservice physical education teacher during a year-long school placement*. [Doctoral thesis, University of Porto].

Moura, A., Graça, A., MacPhail, A., & Batista, P. (2021). Aligning the principles of assessment for learning to learning in physical education: A review of literature. *Physical Education and Sport Pedagogy*, *26*(4), 388–401. doi:10.1080/17408989.2020.1834528

Chapter 13
Using Rubrics as Feedforward Tools for Subject Contextualized Dialogue

Kendall Richards
Edinburgh Napier University, UK

Nick Pilcher
Edinburgh Napier University, UK

ABSTRACT

This chapter outlines an approach to using rubrics in a feedforward process using dialogue between teachers and students that takes place before students hand in their assignments. As such, it aims to complement existing research demonstrating the value of rubrics for feedback and for understanding subject content and assignments. After reviewing some key work on the use of rubrics, the chapter outlines theory around three areas; the importance of context to rubric terminology; the nature of dialogue; the specificity of language to individual assignments to a degree that resists any universal or future transfer to other assignments. The chapter then provides three examples of potential questions to focus dialogue using rubrics to help students understand what is expected of them in their assignments. The approach is then discussed, specifically in terms of the conditions that can facilitate it, and suggestions made for how others could use it.

INTRODUCTION

We work in Higher Education. We are lecturers based in departments. Our roles are to support students (and lecturers) with student work and assignments. One of us is based in a School of Computing. The other is based in The Business School. We teach in modules and alongside subject lecturers. Our remit and goal is to help students understand what is required of them in their academic work here. We are not part of a centralized unit, we do not consider what we do as delivering 'Study Skills' 'Academic Skills' or such like, and feel in fact no such things exist (see Richards & Pilcher, 2020a) a view we fully

acknowledge is not without its detractors nor an uncontroversial one (see Richards & Pilcher, 2021). What we ourselves do is to support students in their subjects, and in the context of the subject, through using the assignments and materials the students are required to know for their subject degrees. Integral to our doing this are rubrics created by lecturers – these provide us with a focus to support students before they hand in their work, and through a process of dialogue in the context of the subject.

Many studies demonstrate the effectiveness of rubrics as tools for feedback (Frey et al, 2018; Pastore & Andrade, 2019; Andrade & Brookhart, 2020), and consider how rubrics work more effectively when given transparently or with teacher dialogue (Wollenschläger et al 2016). Some show individualised feedback to be more effective than generic feedback (Wollenschläger et al 2016) and underline the importance in training students how to interpret feedback (Sadler, 1998). Rubrics are invariably used with the aspiration of helping students approach future tasks through applying what they have learned to these future tasks (Sandrade et al, 2009; Wohlschlaeger et al; 2016).

Nevertheless, some are sceptical such aspirations are achievable (e.g. Sandrade et al., 2009), and others note the unequal power position in the relationship between teachers (understood here to mean teachers, tutors, lecturers etc) and students (Sadler, 1998), noting that teachers have numerous advantages over students in their knowledge of the task, criteria and standards (Sadler, 1998). Indeed, it is stressed that teachers should help students with understanding terminology, and considered essential that teachers "be literate" in the discourses of assessment before they can "help students to recognise (read), realise (speak) and move between the multiple discourses of schooling" (Willis et al, 2013, p. 243). Teachers need to be able to "recontextualise" knowledge to help students understand its function (Bernstein, 1996, p.47)

Such research uses rubrics to help students through a perspective that in essence harnesses them in feed *back*. Undoubtedly, students are provided with rubrics in advance in the form of marking and assessment criteria in guides and handbooks, and told these rubrics will assist them in their work and understanding of what is required. Also undoubtedly, teachers and lecturers will tell students that, for example, an 'excellent level of work will involve or require the aspects that are described in the rubric itself, for example to show 'excellent use of references' or 'excellent illustration of design.' Here the rubric is ostensibly 'fed forward' although only later will students see their 'feedback.' In other words, students are only told how their own assessments align with rubrics *after* their assessment in order to help them to understand their performance (e.g. Sadler, 1998; Sandrade et al., 2009; Wollenschläger et al 2016; Andrade & Brookhart, 2020). Even where rubrics are explicitly stated as being used in a process of feed-forward (De Boer et al., 2021; Hill et al., 2021) the way they are used for feed-*forward* is through a use of them in feed-*back* for *students* to then take these forward. Here, however they are used, the goal is to use rubrics to help students understand learning from a specific individual assessment after it has been done, to then take that understanding forward to future, different, assessments, even if this is named as being feed-forward.

We argue there are a number assumptions involved with this approach; firstly, it is assumed in both cases that lecturer understandings of words and meanings will be shared by students; secondly it is assumed students will know how to interpret the words and meanings in the rubric in a subject context; thirdly, there is an assumption that if students understand the word 'excellent' in one context it applies universally to all others. The latter is both the case where it is assumed feedback from one rubric will help with future assignments, but particularly the case where it is assumed certain terminology and its meaning applies universally. In this chapter we challenge these three assumptions both theoretically and practically and outline an approach where we use rubrics in a feed *forward* process as a focal point for subject contextualised dialogue between teacher and students *before* a specific individual assessment

takes place. Thus, our definition of 'feed-forward' is the use of the rubric as the focal point for dialogue between teachers and learners around a specific assessment *before*, and *not* after, that assessment takes place. Such a process addresses the first and second assumptions that lecturers and students understandings and interpretations of the words will be the same, and that students will understand what the words mean. It also addresses the third assumption that words and their meanings are universally applicable and neutral in meaning and context, both for future assignments and across different subject domains (i.e. that 'high level of criticality' means the same for Design as it does for Engineering).

The remainder of the chapter is as follows: we first provide theoretical underpinning that challenges the first two of the above assumptions through drawing on the work of Mikhail Bakhtin (1981, 1986), Valentin Voloshinov (1929), David Bohm (1996) and Martin Buber (1947) to underline both the importance of context to word meanings, and also of ensuring that dialogue between students and lecturers is both genuine and technical and avoids any danger of becoming 'monologue disguised as dialogue' (Buber, 1947) whereby two individuals engage in a dialogue and leave it believing they have learned something but are in fact no wiser than when they began that dialogue. We then draw on theory from linguists such as Saussure (1959) and their appeal to theory relating to current entrenchments of neoliberal principles and learner focused paradigms in education to illustrate the appeal of universality (Plant, 2011) and responsibilisation (Becker, 1993, 2011) before challenging the above third assumption that language is universally applicable and relevant.

In a subsequent practical section we illustrate, using examples, how we use rubrics to focus lecturer / student dialogue around individual assessments in attempts to avoid falling victim to the above three assumptions. We do this with examples from different subject contexts to illustrate how others can apply the process similarly. We do this with the ultimate goal to improve the standard of student work, learning, and critical abilities in their subjects. In addition we argue it has a number of indirect benefits for lecturers such as reducing a need to set and rewrite assignments where students fare poorly. In a concluding section we draw together the main points from the chapter and suggest avenues for future work.

BACKGROUND:THEORY

Assumptions 1 and 2: Lecturer and Student Understandings of Terms Will Be Similar, and Students Will Understand What Rubric Terms Mean

Mikhail Bakhtin (1981) and Valentin Voloshinov (1927) both underlined the key importance of context to the meaning and interpretation of a word. This context is forever changing, and alters the meanings and interpretations of words over time in a dialogue that is never ending (Bakhtin, 1981). Notably, key aspects such as intonation can severely alter the meaning and interpretation of a word (Voloshinov, 1927) and thus alter how it is understood and interpreted. Words themselves have three owners (Bakhtin, 1986): the addresser; the addressee, and; a neutral, dictionary, meaning. The latter is neutral and removed from context, but when placed in a specific context by addresser and addressee the word assumes its own meaning and value. This value and meaning is itself impacted on by aspects such as intonation, which are inaccessible when the word is only in written form (Voloshinov, 1927). Notably, the dialogue around the meanings and interpretations of words is continual; it is never complete (Bakhtin, 1981). Indeed, should the dialogue ever become complete then the words, according to Bakhtin, assume a sacred qual-

ity whereby they are impervious to questioning (Bakhtin, 1981). This is a theme we visit further when we consider Assumption 3 below.

The key role context plays, and the key role we play in creating that context, are clearly illustrated for us by the comments of a lecturer in clinical psychology who we spoke to on a project where we explored context using physical objects (cited in Richards & Pilcher, 2020b, p.137): "everything exists within context, nothing exists outwith context and we define what the context is and we change the viewpoint according to the context." We also find context to play a key role in other work we have undertaken (e.g. Pilcher & Richards, 2016; 2022) and is something emphasized in other more recent work (e.g. Fecho, 2011).

For us this theory challenges the above two assumptions in both a positive and negative way. What we mean by positive in the specific context here, is the way in which the above theory underlines the importance and key role of exploring and outlining the specific context of the words in a rubric for an individual assignment *before* students undertake that individual assessment. What we mean by negative is how the theory underlines the missed opportunities with *not* exploring and outlining this specific context of the words before the assignment. The positives of outlining the specific context will illustrate to the students how possible rubric words and terminology such as 'critical analysis' apply to a specific context of a specific assignment. They may indeed differ greatly according to whether they are being applied in Design, Nursing, Business, or Computing (Pilcher & Richards, 2016) and by extension other subjects, let alone levels. The dialogue thus must be specific to the subject and use concepts and ideas that are key to the context of that particular subject (Pilcher & Richards, 2016). We have found this ourselves in dialogues we have engaged in (see, e.g. Richards & Pilcher, 2015 for examples) and this is also evidenced in other subject areas (see e.g. Hauk et al, 2014 for examples in mathematics). The subject context is key, as opposed to the context of learning online, or learning face to face, it is the subject context that the students are doing the assignment in that is key and for this reason we argue that centralized support will by its very nature be decontextualized (Richards & Pilcher 2020a). Providing this context of anchoring the discussion to the specific subject and its assignment contextualises the words to what the students need to do. From our own studies we know that a 'report' and 'essay' differ from one subject to another (Richards & Pilcher, 2019) and that different subjects have different understandings of the meanings of words according to key ideological tenets (for example where 'empathy' means understanding patient feelings to a nurse but means alignment with the client to a designer (Richards & Pilcher, 2016)).

Here then, outlining and exploring such words in their specific subject assignment contexts using dialogue around rubrics and before students undertake individual assignments outlines to them how they are supposed to complete the work and what is expected of them. It also has a number of additional positives we argue, such as it being integral to learning of material, engaging students more fully in the process of assessment, helping reduce stress, and in addition subsequently avoiding any negatives by reducing the need for subsequent work in arranging, writing and marking resit assessments for students who fail. In addition, it has the positive impact in that it aids managing student expectations of grades, work, and abilities, and avoids the perennial post assignment question of 'Why didn't you tell me that before?'

A key question is of course, how should this dialogue take place? Theoretically, Bohm (1996) writes of the need to suspend assumptions in order for any dialogue to take place. For us, this is closely related to the need to avoid thinking that any meanings of words are fixed, solid, possibly sacred, and impervious to questioning (Bakhtin, 1981). For Bohm, an indication that assumptions need to be suspended is conveyed through feelings of frustration, anger, or confusion. For example, if dialogue is taking place around the word 'Essay' and the teacher is conveying their own understandings of what an essay is to

students and it differs greatly from what other teachers (or indeed generic glossaries and guides) have said, it is possible students may feel confused or frustrated. Here, the key would be to inform them that in order to understand what was required for the specific essay and for this teacher it will be necessary to let go of this confusion or frustration and explore what is being discussed openly. Only in this way can an open dialogue take place (Bohm, 1996). In terms of how such open dialogue can be initiated, we suggest an opening recommendation from the teacher can help here. This could be along the lines of something like 'You may have done essays for others before, but the term 'essay' can mean many different things, for me it means…..' (see practical section below).

Regarding types of dialogues, Buber (1947) wrote of three different types of dialogue: genuine, where the dialogue focuses on considering people as fellow human beings; technical, where the dialogue is focused on a specific subject, and; monologue disguised as dialogue, whereby two individuals leave a dialogue feeling they have learnt something, but in fact are no wiser than when they started. For us, an example of 'monologue disguised as dialogue' would be whereby a student was to ask a lecturer a question along the lines of 'Is Discuss where you want us to consider one side against the other?' and the lecturer answers 'Yes' (Richards & Pilcher, 2015). Here, we consider that the student will leave the dialogue feeling they know what 'Discuss' means, and the lecturer will leave the dialogue feeling they are happy the student knows what to do. However, we argue that in order for this to be the case the dialogue needs to be more exploratory and genuine and technical in nature. What we mean by this is, for example, when the student asks 'Is Discuss where you want us to consider one side against the other?' rather than reply simply with 'Yes', the lecturer would explain more by replying with something along the lines of 'Yes, but 'Discuss' for this assignment has particular aspects it wants you to consider, what do you think these could be?' to enable further dialogue (see the practical section below for more details). Here, the dialogue will be genuine in that it is focused on students as human beings, but it is also technical in that it focuses on the specific subject at hand that the assignment is set for.

Before turning to outlining the theory we feel counters Assumption 3, we want to consider a fear or a concern that many teachers we have spoken to have about using this approach. This concern is closely related to the idea commonly conveyed as: 'Yes, I get all this, but I want to see what the students can do, I don't want to give them the answer'. We outline this here as we can see this as a concern or frustration arising and we wish to focus on this now to outline how we see a dispelling of this assumption as helping continue the dialogue here (Bohm, 1996). What we commonly stress is that this process does not actually give students the answer; in fact, it gives the lecturers more control over what the students do. It still retains the key aspect of assessment that it is the students own work that is being assessed. It does this because the dialogue is supported and illustrated by examples selected by the teacher themselves, so the teacher is in control of what to tell students and what not to tell students. In other words, and as we ourselves often find that we say, the phrase 'I don't want to give you the answer, I want to see what you can do' is one that we often iterate to students. We now consider theory that counters Assumption 3.

Assumption 3: Language From Rubrics Applies Universally—Both for Future Assignments Students Need To Do and Across Different Subjects

The 'father of linguistics' is considered to be Ferdinand de Saussure, who is said to have ushered in a revolution of Copernican proportions to the field (Harris, 2013). Saussure's revolutionary insight was that language is anything but fixed, and that a word does not represent the connection between an item and its reality, rather, a word represents a psychological bond in the head of the language user between

the signifier of the 'sound-image' made by a speaker, and what is signified by it. This connection, or bond, is purely arbitrary, and will differ depending on whether the language user is in one part of the world using one language or in another part of the world using another (Saussure, 1959). Nevertheless, where this arbitrary quality of language stopped for Saussure was when the individual speaker had the bonds of the language 'fixed' in their heads. For Saussure, once these bonds were 'fixed' the language did not change for the user. What this quality or facet meant for Saussure was that language was both 'synchronic', in that for the user once fixed it never changed, but at the same time it was also as a system 'diachronic' and changed over time. For Saussure, this meant that it could be acknowledged that language changes slowly over time (over a generation commonly) but that for the individual speaker it remained fixed and was constitutive of a system. In turn, this system was external to the individual, and could be taken away for the purposes of analysis and pedagogy. In Saussure's words: "language stands out as a well-defined entity… external to the individual, who by himself is powerless either to create it, or to modify it" (Saussure, 1959, p.31). Not only this, but notably "language, unlike speaking, is something we can study separately…. Whereas speech is heterogeneous, language, as defined, is homogenous. It is a system of signs in which the only essential thing is the union of meanings and sound-images, and in which both parts of the sign are psychological… language is concrete" (Saussure, 1959., p.15). Thus, language had a structure, was a system, was eternal to the speaker, remained relatively constant over time, and although arbitrary, once fixed in the head of a user, did not change. If this was the case then language could be studied, taught, and assumed to be both constant and universal.

Such theory was critiqued by the work of Valentin Voloshinov (1929) of the Bakhtin school although this critique was not translated into English until the 1970s. For Voloshinov, Saussure's view of language was an abstract objectivist one: the language could be abstracted from its context for study, and was objectivist and constant. For Voloshinov, this ignored the key role played by context and ran counter to his own theory that language was created by people and was individual and subjective to that individual (1929). Any written text for Voloshinov only represented what he termed as being the 'inert hardened crust' of language activity that was unable to reveal its active creative nature and true context (ibid, 1929). For Voloshinov, rather than being fixed, language was constantly changing and interpreted subjectively by individuals, and used artistically.

In studies we have done, we have found that individuals use what appear to be the same words (cf Saussure, 1959) but that their understandings and interpretations of them are unique to subject context and individual (cf. Voloshinov, 1929). Thus, whilst on the surface the language may look the same and give a surface appearance of objectivity, underneath this surface it is highly individual and subjective (Pilcher & Richards; 2016; 2017; 2022; Richards & Pilcher, 2014; 2015; 2019; 2020b). In other words, it may look, and appear, as though the language that students use in one rubric is the same as the language they will need for other rubrics and that it applies universally, but in fact, once this language is explored in individual contexts it is highly varied and interpretations of it are highly context specific. There is therefore no 'magic bullet' of teaching students what terms mean as the terms from one assignment and its context will differ in their meaning in a future one.

In terms of their contexts and times of writing, Saussure was writing in the late nineteenth and early twentieth century, and Voloshinov in the Soviet Union in the early twentieth century. Saussure was very much writing from a perspective linked with psychology, whereas Voloshinov was much more grounded in a focus on literature. It is understandable therefore that one would seek to find a structure in language that linked to the psychology of individuals, and that the other would seek to find interpretations from a literary perspective. Nevertheless, the work of Saussure set the study of language and linguistics along

a structural route that has sought to identify structures of genres of language use (e.g. Swales, 1990) and to identify key recurrent words for students to learn that are structurally grounded (e.g. Coxhead, 2000). In contrast, the work of Voloshinov has been more in the literary vein such as that of William Empson (e.g. Seven Types of Ambiguity, (1930)) and when attempts have been made to add to the field of linguistics through the use of such theory it has often fallen on deaf ears. Empson himself found this when his work 'The Structure of Complex Words' (1951) was completely ignored by the linguistics field, and we have also found this in our own work (Richards & Pilcher 2016; 2017). Thus, whilst of their time, and having been developed and critiqued, both Saussure and Voloshinov have nevertheless informed how the field of linguistics and literature has developed, although any dialogue has often followed the parameters they themselves outlined and not often crossed fields.

Arguably, the current paradigm of viewing education and learning is locked into what Barr and Tagg (1995) outlined as the 'learner-focused' paradigm. This was a shift in the 1990s from the teacher to the learner, and, whilst the teacher has remained continually valued (e.g. Sadler, 1998; Willis et al., 2013) it is arguable that the focus on the use of rubrics in the feed *back* capacity is grounded in a paradigm that focuses on the learner and on attempting to empower the learner to move forward themselves. This can be argued to be the case not least because such research and such usage of rubrics is intended to provide the 'magic bullet' of harnessing learners with the knowledge to help themselves, in line with the indomitable encroachment of neoliberalism (Olsen & Peters, 2005) and its key tenets of responsibilisation (Bonnano, 2017), development of human capital (Becker, 1993, 2011) and universality of human abilities and transferability (Peck, 2010, Plant, 2011).

So entrenched is neoliberalism currently that we fear it has led to a situation whereby its tenets and key facets have become sacred and are above dialogue and questioning (Bakhtin, 1981). Learners are customers, centralised services are based on language being universal (Richards & Pilcher, 2020a), centralised glossaries explain what terms mean, the language must therefore mean the same thing and apply universally (cf. Saussure, 1959; Plant, 2011). Learners must have the responsibility themselves (cf. Becker, 1993, 2011) to learn a language and its terms through consulting sources such as dictionaries and centralised glossaries based on a theory that the language can be abstracted and is objective (Saussure, 1959). Such theories we fear may have created whole systems and structures that are continually reinforced by people's actions (Giddens, 1984) to the extent that they are now sacred and impervious to any dialogue (Bakhtin, 1981).

We now outline a practical approach that is based on challenging these assumptions with the aim for others to reopen dialogue around the terms used in rubrics before students do their assessments. We do this by first focusing on the context of rubrics and then outlining three practical examples from our use of the feed-forward process in dialogue with student groups.

MAIN FOCUS OF THE CHAPTER: RUBRICS AS FEED-FORWARD TOOLS IN FOCUSED DIALOGUE PRIOR TO ASSESSMENT DEADLINES

Context

A key discussion point around rubrics urges that debate needs to move beyond whether rubrics are good or bad in themselves and to consider instead how best to use them (Turley & Gallagher, 2008). In work that focuses on the use of rubrics in specific subject contexts often tends to use rubrics for learning

processes after the students have completed their assessments. In a Chemical Engineering context, for example Cifriuan et al (2020) have used rubrics post assessment and have encouraged student participation in co-evaluation of work and rubric creation (cf. Becker, 1993; 2011). Indeed, Cifrian et al (2020) talk about how they use rubrics for reflection on performance with students in very open group discussions that involve dialogue between teachers and students: "During these face-to-face group classes teachers answer any queries, open dialogue with students to interpret rubrics and understand how they are manifested in the project report" (Cifrian et al., 2020, p65). However, the specific nature of this dialogue is not outlined, and it is again using rubrics post-assessment submission, indicated by the fact that rubrics here are used for critical analysis and reflection, as rubrics are used in the classes "with the object of stimulating students' capacity for critical analysis and personal reflection" (ibid, p.65). The use of dialogue around rubrics post-assessment is also mentioned in other work even if it may be defined as 'feed-forward' (Hill et al., 2021). With work that states it uses rubrics as teaching tools teachers are not employed as guides for a dialogue with students about them, rather they are used by students in group discussions, for example in Calculus (Auxtero & Callaman, 2020) or in educational methodology (Zhang et al., 2019), or with Sports students (Bradley et al., 2020). In other work, rubrics have been used to help instructors integrate subject content, for example in bioinformatics and computing (Tapprich et al., 2021)

Similarly, in science education students co-created their own rubrics which were then used to assess their work (Luft, 1999) but no specific connection between rubrics and examples from the teacher are outlined. In a biological science context, somewhat similarly it is aspired that rubrics help students to self-direct their own learning (cf. Becker, 1993, 2011) and help teachers develop and refine assessments (Allen & Tanner, 2006) sometimes specifically targeted at University level (Velasco-Martinez & Tojar-Hurtado, 2018). In Engineering as well rubrics have been used after students complete assignments to show them how their work has been graded in terms of its quality (Wright et al., 2022). Key to all these uses of rubrics is the theme of using rubrics to responsibilise students (cf. Bonnano, 2017), encapsulated by De Boer et al (2021) who note that when rubrics are used in feed-back and discussed with a view to their contents and the feed-back later being used as feed-forward so to help students "define their own learning goals and *take responsibility* to develop on these [our italics]" (De Boer et al., 2021, p.42).

Thus, where feed-forward is mentioned it is as a part of how the feedback can help with future work (De Boer et al., 2021) and where dialogue is explicitly stated as occurring it is in a context of the use of rubrics in discussion after the assessment has taken place (Cifrian et al., 2020) even if it is described as feed-forward' (Hill et al., 2021). Notably, De Boer et al (2021) comment on how "a rubric alone may not be enough for students to recognize and understand this thing that is new to them. They will need guidance and support to become familiar with the concepts and context" (De Boer et al., 2021, p.22). We now outline how we have used rubrics in a feed-forward process with students before a specific assignment deadline so the rubric is used to teach students what is expected of them before they do their assignment.

Practice

We outline three examples from our own use of rubrics in focused dialogue in a feed-forward process prior to the assessments taking place: one from a first year computing module; one from a first year module in the subject area of critical literature and; one from an MBA module where we plan to use the feed-forward dialogue based approach. All rubrics are in the appendix and we draw on examples of them here. We outline possible questions we could ask after briefly outlining each assessment. Whilst we fully acknowledge that there are power imbalances between students and teachers and students may

not wish to speak for fear of these power imbalances, we note that we explicitly emphasize this in the tutorials and classes in an attempt to make sure all dialogue is public and there are no 'hidden transcripts' (Scott, 1990) of things left unsaid.

Example 1: First Year Computing Module

The task for this first year module that the students do is as follows: "Write a 1500 word (+ or – 10%) case study report responding to the following: **"Select an organisation, preferably one to which you have personal access/knowledge. Analyse the environment, structure and culture of the organisation and focus on one system/app/technology/program relevant to your degree programme that currently or potentially benefits the organisation."** This should be supported by relevant references from academic journals, books, and real-life observation or interviews, and company information (but not Wikipedia)."

In this rubric (see the Appendix for an extract) for this assessment the criteria are down the side on the left, the grade categories (from 10 (highest) to 0 (lowest) is along the top, and the rubric descriptions for each criteria are in the boxes below. If we take Criterion 5: '**Criterion 5 Range, currency and appropriateness of your sources and references; Range, currency and appropriateness of your sources and references. Consistency and formatting of in-text citations and reference list.'** Category 10 (the highest mark) is as follows: *"Excellent range, currency and relevance from research and reputable sources, ample to provide evidence for your points. Fully accurate citations, citations tally with bibliographical entries, and bibliography is in the correct APA format. Cannot be improved."*

Prior to our engaging in dialogue with the students for this rubric, they will have been talked to by a subject specialist librarian who will have pointed them in the direction of key sources for the assignment. Students will then come to the next group tutorial with some sources that they have selected on the basis of this talk and it is these sources that then form the start of the dialogue. In addition, we tell them to bring a peer reviewed source and any other non-academic source.

We would have, and have here examples of questions and possible responses we have experienced in square brackets afterwards to focus this dialogue along the lines of:

- "So, if you want to get a mark of 10 it says you need to provide 'Excellent range, currency and relevance from research and reputable sources' so what is an 'Excellent' source do you think?"
- "What is currency for the sources in this specific context for computing?"
- "What is a reputable source?"
- "So the subject librarian talked about things like peer reviewed sources, date of publication, based in the subject and using computing databases….. so – what does that all mean for this report in this subject?"
- "So I can see you have brought in this source "Tesco use of IT and information systems – introduction to IS" and it is a Student administered Case Study – so let's look at this one in the context of the report"
- "When was it published? Well, it says 2009, does that matter in terms of the technology? What has developed since then? So is it out of date?" [yes]
- "So how could you use it?" [As a starting point, historical context]
- "Yes, that would work – now tell me where was it published and where is the study based?" [published by Dr Phil Kelly – focusing on the UK]

- "Great – so would you mention all this in the case study you write?" [yes]
- "One other thing you must do, is you have to say how the source is relevant to the task, how would you do that with this one?" [by highlighting it is about a company and their use of IT]
- "Absolutely, and mentioning all this is you demonstrating you have an 'excellent' understanding of the answer"
- Thinking specifically of the rubric here with these questions, what then would constitute an 'Excellent range, currency and relevance'?
- What in a computing context are 'reputable sources'?

Example 2: First Year Module in Literature and Adaptation

The task for this module is an essay (1,300 words) that students write based on their interpretation of how a character's Vampire status functions symbolically in a Graphic Novel

Option B1. A Girl Walks Home Alone at Night Graphic Narrative essay:

Question: How does The Girl's vampire status function symbolically in the text?

Choose up to two pages of the comic (or six panels in total from across single or multiple issues) and analyse them to support your argument. Quote and analyse the text and the images. You **must** draw on relevant academic sources (a minimum of two sources) such as the essays and chapters digitized on Moodle that deal with theories and issues in adaptation studies, graphic narratives and/or genre writing."

This is not a subject we are familiar with, certainly not as much as the students, and our dialogue around this one is targeted towards reaching a mutual shared understanding. An extract for the rubric for this is in the Appendix and is what we see as being highly decontextualized and generic, and consequently our dialogue focuses on contextualizing and specifying what is meant by the rubrics.

The rubric students work form here goes for all categories and is subdivided according to classification of degree: '80+ First class'; '70+ First class' '60+ Upper second class' etc – if we take as an example some details from the highest and the lowest categories so:

'80+ First class …..Very thorough, detailed analysis with highly convincing, wide-ranging evidence to back up arguments' and the same category from the much lower category of '30+ Fail (but with potential for compensation)…. The work consists of a string of assertions and opinions which may not related to each other, with little or no use of supporting evidence'

As examples, we would have questions to focus this dialogue with the group along the lines of:

- "So the question is 'How does The Girl's vampire status function symbolically in the text?' of his graphic novel – What does 'symbolically' actually mean?"
- "What are the main theories in adaptation studies?"
- "What has it been an adaptation from? How do they do this to ensure the adaptation to this genre and style?"
- "Can you give me some examples from the graphic novel here of how these theories show the vampire functioning symbolically?"
- "What sources would you use to justify your choice of example here? And How?"
- "How do you use the source to 'back up' your argument? How would it be only a 'strong of assertions?"

- "One of the recommended theory texts is by Julie Sanders on Adaptation. How would you integrate this into your answer to 'back-up' your argument?' How could you mention the theory but it would only be an assertion?"
- Thinking again specifically of the rubric here, what, based on the above would be 'highly convincing, wide-ranging evidence'?
- What would be a Fail? What would be an example of a 'string of assertions and opinions'?
- What would constitute 'little or no use of supporting evidence'?

Example 3: Postgraduate Module for a Masters of Business Administration

The assessment for this module is as follows: "A patchwork text (a specific type of portfolio) in which you will present evidence of your learning and development generated during the module submitted with a critically evaluative account and an action plan (2,500 words)." Students are given suggestions on how to structure the text and on which reflective framework to use. An extract for this Rubric for this assessment is in the Appendix.

We would have, and have had examples of the following questions to focus this dialogue with the tutorial groups along the lines of:

- "So, it's asking you to use Gibbs as a framework, so would you use that?"
- "When I say 'would' do I mean it's possible to use it or you should use it?"
- "Can you refer to other reflective frameworks?"
- "What would an 'excellent' outline of the rest of the report be?"
- "Would such an outline include all parts?"
- "What would a 'poor outline be?"
- "Section A wants you to focus on 'Task Focus; Academic Expectations; MBA Content; what was new or unexpected' – how will you do this?"
- "What would be evidence of these things?"
- What is 'MBA Content' that is asked for in the rubric?
- How exactly, as the rubric suggests, would you use Gibbs here?

DISCUSSION

As noted, above, one common fear or caution expressed by many who we explain our approach to is along the lines of 'Yes, but I want to see what they can do, I don't want to tell them the answer.' We hope the above examples show that in our dialogues we do not give students the answers, rather, we focus them on what it is the assignment wants them to do. We focus dialogue on exploring the nature of the content rather than the content itself. The process is done from the perspective of someone learning about the assignments and on the basis that the subject context (Bakhtin; 1981; 1986; Voloshinov; 1927;1929) is what is key, and that what would be 'excellent' in a first year Computer Science assignment will be different to what is 'excellent' in a postgraduate MBA assignment (cf. Bakhtin, 1986). What such dialogue aims to do is to align student (addressee) understandings with those of the teachers (addressers) and to do so by contextualizing otherwise neutral and dictionary definition type terms into their subject contexts. Here the process is directly aligned with learning and with improving student knowledge of the

subject through focusing the dialogue in the subject context and on the areas that students are expected to produce work in. By focusing the dialogue on the rubric, this has directly shown students what is expected of them, helped clear up any misunderstandings they have had, and helped ensure they learn better. It has, in short, very much helped avoid the question later down the line of 'Well, why didn't you tell me that earlier?'

Notably, by doing this *before* the assignment is also an explicit recognition that work with one rubric is not a 'magic bullet' for all other, future assignments, as it is implicitly grounded in the idea that all assignments are unique and all different. This is also why we do not group together questions for others in a table – we cannot do this as all assignments are different and unique. Nevertheless, we argue that the three examples we outlined above provide key indicators for how others can focus dialogue around their own rubrics to help students learn and to avoid students from producing work which is off-topic. Indeed, we argue, however, that making students aware of this fact is a helpful one, and we hope (and certainly tell students) to ask similar questions of their other subject lecturers that we outline above. Undoubtedly students can therefore take the process forward of engaging in dialogue, but the approach is grounded in the key paradigm that rubrics in the form of feedback can only give information on the assignment they are attached to, and that this will not necessarily feed forward into future assignments. Philosophically, we would not expect feedback on an assignment in one module on HRM to feed forward into a module on Financial Accounting. Students may be able to take some lessons forward, but the assignment will inevitably be different.

As outlined above, in terms of the nature of the dialogue we engage in, the aim is always to try and steer it towards technical dialogue (Buber, 1947), to avoid falling into any traps of engaging in monologue disguised as dialogue, and to suspend assumptions (Bohm, 1996) about the nature of assignment tasks for specific assignments, and how they may or may not be similar to other tasks. Our aim with the examples is to demonstrate to others how to focus and fashion dialogue around rubrics in this feed forward approach before the assignment is submitted. Here, our aim is to complement existing studies using rubrics in feedback (e.g. Frey et al, 2018; Pastore & Andrade, 2019; Andrade & Brookhart, 2020) and also those that use rubrics to feedforward through feedback (De Boer et al., 2021; Hill et al., 2021) and help build on the work that underlines the key role rubrics can play in learning and assessment (e.g. Sadler, 1998; Sandrade et al., 2009; Wollenschläger et al 2016; Andrade & Brookhart, 2020). We also wonder that such approaches could be used in collaborative rubric design between teachers and students (cf. Kilgour et al., 2020) to help make rubrics clearer and more focused towards helping students produce the type of work that teachers both value and strive to help them do so in their education.

Who can do what we ourselves do? Or, perhaps more appropriately, what conditions are necessary for what we suggest here to take place? Context is key, and by context we mean subject context, so a key condition is that dialogue takes place in the specific subject assignment and rubric that students are tasked to do. Subject knowledge is also key, and this can come from either the students, but, critically, also needs to come from subject lecturers, or those who can communicate with the subject lecturers to ask about content. It is in the latter category that we operate, and when we do so the subject lecturers are either in the room or contactable via email to ask about the assignments and relay this information to students. Indeed, we often find ourselves commenting to students along the lines of 'That's a very good question, I'll check with X and get back to you'. Conversely, what this means is that what we suggest here cannot be done in conditions of 'generic' workshops as the context is key, and the subject knowledge of the specific assignment is key. To try and do what we suggest here outside these conditions is not possible, and is based on a false assumption that all terminology in rubrics is universally applicable.

Nevertheless, in the right conditions, and by those either who are subject experts or who can find out about the subject, we have found the above process to help students succeed, to have indirect benefits of then saving others time through a reduced necessity to file Academic Conduct Cases or to create resits, and, ultimately, to help students learn through an approach that questions the learner focused paradigm and draws on teacher knowledge.

CONCLUSION

In this Chapter we have outlined the underpinning theory and given practical examples of how we ourselves engage in dialogue with students through the use of rubrics in a feedforward process which takes place before students submit their assignments. Theoretically we have presented a number of concepts from key thinkers (Bakhtin; Voloshinov) that underline the key role played by context and the importance of the subject context (our own studies) to the specific interpretation of terminology and assessment forms. We have also outlined theory around the nature and process of dialogue to try and focus dialogue to be technical and to avoid lapsing into what Buber (1947) described as 'monologue disguised as dialogue' whereby two individuals may feel they have learned something but leave a dialogue no wiser than when they entered it. In addition, we outlined linguistic theories that see language as being something that can be abstracted from its context for study and is objective; how these theories align perfectly with the currently entrenched neoliberal political philosophies, but how we feel these theories do not reflect the reality of what students are faced with or the nature of the work they are required to do.

We then outlined examples of how we draw on rubrics to engage in dialogue in our roles as lecturers who support students to help students understand what is required of them in a feedforward process that takes place before students complete their assignments. We noted that essential conditions for such dialogue are that it is done in the specific context of the subject assignment and by those with subject specialist knowledge, or with easy access to those with such knowledge. Who can do this? Anyone who can work in these conditions; teachers; others in academic support who are able to support students in conditions outlined here. We stressed that the process of doing this before the assignment is not telling students any answers, rather, we feel it is illustrating to the students what is required of them of each specific assignment, as well as outlining to them how future assignments may differ, and how they can ask other teachers about what they need to do to succeed in them. We also outlined how we feel this process can indirectly help teachers through reducing a need for teachers to set reassignments or to file cases of Academic Conduct. Whilst on the one hand our approach could be said, somewhat counterintuitively, to align with many neoliberal concepts such as responsibilisation (empowering students to ask); marketization (improving retention, word of mouth feedback; rankings) and with helping ameliorate the pressures of massification (by helping more students do well and avoiding a need for resits); it can, in our experience, be best done with smaller groups of students, although the limits of this is not one we have tested and could perhaps be something future research could consider.

REFERENCES

Allen, D., & Tanner, K. (2006). Rubrics: Tools for making learning goals and evaluation criteria explicit for both teachers and learners. *CBE Life Sciences Education*, *5*(3), 197–203. doi:10.1187/cbe.06-06-0168 PMID:17012210

Andrade, H. L., & Brookhart, S. M. (2020). Classroom assessment as the co-regulation of learning. *Assessment in Education: Principles, Policy & Practice*, *27*(4), 350–372. doi:10.1080/0969594X.2019.1571992

Andrade, H. L., Wang, X., Du, Y., & Akawi, R. L. (2009). Rubric-referenced self-assessment and self-efficacy for writing. *The Journal of Educational Research*, *102*(4), 287–302. doi:10.3200/JOER.102.4.287-302

Auxtero, L. C., & Callaman, R. A. (2020). Rubric as a learning tool in teaching application of derivatives in basic calculus. [Journal of Research and Advances in Mathematics Education]. *JRAMathEdu*, *6*(1), 46–58. doi:10.23917/jramathedu.v6i1.11449

Bakhtin, M. M., & Holquist, M. (1981). *The dialogic imagination: Four essays*. University of Texas Press.

Bakhtin, M. M., Holquist, M., McGee, V., & Emerson, C. (1986). *Speech genres and other late essays*. University of Texas Press.

Barr, R. B., & Tagg, J. (1995). "From Teaching to Learning—A New Paradigm For Undergraduate Education." Change. *Change*, *27*(6), 12–26. doi:10.1080/00091383.1995.10544672

Becker, G. S. (1993). *Human Capital*. University of Chicago Press. doi:10.7208/chicago/9780226041223.001.0001

Becker, G. S. (2011). Foreword. In *The Oxford Handbook of Human Capital, edited by A. BurtonJones and J.-C. Spender, ppxiii–ppxxvi*. Oxford University Press. doi:10.2307/j.ctt6wq7jd.3

Bernstein, B. (1996). *Pedagogy, symbolic control and identity: Theory, research and critique*. Taylor and Francis.

Bohm, D. (1996). *On dialogue*. Routledge.

Bonanno, A. (2017). *The Legitimation Crisis of Neoliberalism: The State, Will-Formation, and Resistance*. Springer. doi:10.1057/978-1-137-59246-0

Bradley, E., Anderson, S., & Eagle, L. A. (2020). Use of a marking rubric and self-assessment to provide feedforward to level 5 undergraduate Sport students: student perceptions, performance and marking efficiency. *Journal of Learning Development in Higher Education, 18*.

Buber, M. (1947). *Between man and man* (R. G. Smith, Trans.). Routledge & Kegan Paul.

Cifrian, E., Andrés, A., Galán, B., & Viguri, J. R. (2020). Integration of different assessment approaches: Application to a project-based learning engineering course. *Education for Chemical Engineers*, *31*, 62–75. doi:10.1016/j.ece.2020.04.006

Coxhead, A. (2000). A new academic word list. *TESOL Quarterly*, *34*(2), 213–238. doi:10.2307/3587951

de Boer, I., de Vegt, F., Pluk, H., & Latijnhouwers, M. (2021). *Rubrics-a Tool for Feedback and Assessment Viewed from Different Perspectives: Enhancing Learning and Assessment Quality*. Springer. doi:10.1007/978-3-030-86848-2

Empson, W. (1930). *Seven types of ambiguity*. Chatto and Windus.

Empson, W. (1951). *The Structure of Complex Words*. Chatto & Windus.

Fecho, B. (2011). *Teaching for the students: Habits of heart, mind, and practice in the engaged classroom*. Teachers College Press.

Frey, N., Fisher, D., & Hattie, J. (2018). Developing "assessment capable" learners. *Educational Leadership*, *75*(5), 46–51.

Hauk, S., Toney, A., Jackson, B., Nair, R., & Tsay, J. J. (2014). Developing a Model of Pedagogical Content Knowledge for Secondary and Post-Secondary Mathematics Instruction. *Dialogic Pedagogy, 2*.

Hill, J., Berlin, K., Choate, J., Cravens-Brown, L., McKendrick-Calder, L., & Smith, S. (2021). Can Relational Feed-Forward Enhance Students' Cognitive and Affective Responses to Assessment? *Teaching & Learning Inquiry*, *9*(2), n2. doi:10.20343/teachlearninqu.9.2.18

Kilgour, P., Northcote, M., Williams, A., & Kilgour, A. (2020). A plan for the co-construction and collaborative use of rubrics for student learning. *Assessment & Evaluation in Higher Education*, *45*(1), 140–153. doi:10.1080/02602938.2019.1614523

Luft, J. A. (1999). Rubrics: Design and use in science teacher education. *Journal of Science Teacher Education*, *10*(2), 107–121. doi:10.1023/A:1009471931127

Olssen, M., & Peters, M. A. (2005). Neoliberalism, Higher Education and the Knowledge Economy: From the Free Market to Knowledge Capitalism. *Journal of Education Policy*, *20*(3), 313–345. doi:10.1080/02680930500108718

Pastore, S., & Andrade, H. L. (2019). Teacher assessment literacy: A three-dimensional model. *Teaching and Teacher Education*, *84*, 128–138. doi:10.1016/j.tate.2019.05.003

Pilcher, N., & Richards, K. (2016). The paradigmatic hearts of subjects which their 'English' flows through. *Higher Education Research & Development*, *35*(5), 997–1010. doi:10.1080/07294360.2016.1138455

Pilcher, N., & Richards, K. (2017). Challenging the power invested in the International English Language Testing System (IELTS): Why determining 'English' preparedness needs to be undertaken within the subject context. *Power and Education*, *9*(1), 3–17. doi:10.1177/1757743817691995

Pilcher, N., & Richards, K. (2022). *Enhancing Student Support in Higher Education: A Subject-focused Approach*. Springer Nature. doi:10.1007/978-3-030-81724-4

Richards, K., & Pilcher, N. (2014). Contextualising higher education assessment task words with an 'anti-glossary' approach. *International Journal of Qualitative Studies in Education : QSE*, *27*(5), 604–625. doi:10.1080/09518398.2013.805443

Richards, K., & Pilcher, N. (2015). Avoiding dialogues of non-discovery through promoting dialogues of discovery. *Dialogic Pedagogy, 3*.

Richards, K., & Pilcher, N. (2016). An individual subjectivist critique of the use of corpus linguistics to inform pedagogical materials. *Dialogic Pedagogy: An International Online Journal, 4*.

Richards, K., & Pilcher, N. (2017). Should we teach from materials developed with corpus linguistics? *English for Specific Purposes Special Interest Group Journal*.

Richards, K., & Pilcher, N. (2019). How a view of language underpins approaches to supporting higher education students that facilitate neo-liberalism, and how to resist this. *Power and Education*, *11*(1), 51–68.

Richards, K., & Pilcher, N. (2020a). Study Skills: Neoliberalism's perfect Tinkerbell. *Teaching in Higher Education*, 1–17.

Richards, K., & Pilcher, N. (2020b). Using physical objects as a portal to reveal academic subject identity and thought. *Qualitative Report*, *25*(1). Advance online publication. doi:10.46743/2160-3715/2020.4023

Richards, K., & Pilcher, N. (2021). *Study Skills are not the answer to students woes*. WonkHE. https://wonkhe.com/blogs/study-skills-are-not-the-answer-to-students-academic-woes/ Last Accessed January 2023

Sadler, D. R. (1998). Formative assessment: Revisiting the territory. *Assessment in Education: Principles, Policy & Practice*, *5*(1), 77–84. doi:10.1080/0969595980050104

Scott, J. C. (1990). *Domination and the arts of resistance: Hidden transcripts*. Yale university press.

Swales, J. (1990). *'Genre Analysis: English in Academic and Research Settings*. Cambridge University Press.

Tapprich, W. E., Reichart, L., Simon, D. M., Duncan, G., McClung, W., Grandgenett, N., & Pauley, M. A. (2021). An instructional definition and assessment rubric for bioinformatics instruction. *Biochemistry and Molecular Biology Education*, *49*(1), 38–45. doi:10.1002/bmb.21361 PMID:32744803

Turley, E. D., & Gallagher, C. W. (2008). On the" uses" of rubrics: Reframing the great rubric debate. *English Journal*, 87–92.

Velasco-Martínez, L. C., & Tójar-Hurtado, J. C. (2018). Competency-Based Evaluation in Higher Education—Design and Use of Competence Rubrics by University Educators. *International Education Studies*, *11*(2), 118–132. doi:10.5539/ies.v11n2p118

Voloshinov, V. N., Matejka, L., & Titunik, I. R. (1973). *Marxism and the philosophy of language*. Seminar Press. [original 1929]

Voloshinov, V. N., & Titurin, I. R. (1987). Freudianism: A Critical Sketch. New York: Seminar Press [original 1927] In Morris, P (Ed.) The Bakhtin Reader: Selected Writings of Bakhtin, Medvedev, Voloshinov.

Willis, J., Adie, L., & Klenowski, V. (2013). Conceptualizing teachers' assessment literacies in an era of curriculum and assessment reform. *Australian Educational Researcher, 40*(2), 241e256

Wollenschläger, M., Hattie, J., Machts, N., Möller, J., & Harms, U. (2016). What makes rubrics effective in teacher-feedback? Transparency of learning goals is not enough. *Contemporary Educational Psychology*, *44*, 1–11. doi:10.1016/j.cedpsych.2015.11.003

Wright, K., Slaboch, P. E., & Jamshidi, R. (2022). Technical writing improvements through engineering lab courses. *International Journal of Mechanical Engineering Education*, *50*(1), 120–134. doi:10.1177/0306419020939621

Zhang, Y., Chen, B. L., Ge, J., Hung, C. Y., & Mei, L. (2019). When is the best time to use rubrics in flipped learning? A study on students' learning achievement, metacognitive awareness, and cognitive load. *Interactive Learning Environments*, *27*(8), 1207–1221. doi:10.1080/10494820.2018.1553187

APPENDIX: EXTRACTS AND EXAMPLES FROM RUBRICS REFERRED TO ABOVE

Example 1: Extract only: There are five criteria and the range from 10 to 0 includes all numbers

	10	8	3	1	0
Criterion 5 Range, currency and appropriateness of your sources and references; Range, currency and appropriateness of your sources and references. consistency and formatting of in-text citations and reference list.	Excellent range, currency and relevance from research and reputable sources, ample to provide evidence for your points. Fully accurate citations, citations tally with bibliographical entries, and bibliography is in the correct APA format. Cannot be improved.	Very good range, currency and relevance, from reputable sources. Fully sufficient to support your points. Almost completely accurate citations, tally with bibliographical entries. Bibliography is in the correct APA format.	Too few references included, or not in APA, and omissions where a reference was necessary but not provided. Mainly unreliable sources from the web rather than from reputable sources such as journals and books. Please refer to the School of Computing referencing guide available via Moodle.	No references, or from unreliable sources from the web or inappropriate to the subject matter. Correct referencing has not been followed. Please refer to the School of Computing referencing guide available via Moodle.	Missing

Example 2: Extract: There are more classification categories but only the first class is copied here

Assessment criteria for essays
80+ First class

- Research has been exceptionally thorough and wide-ranging, information has been selected and presented very appropriately for the topic, and the writer has contributed original ideas to the discussion
- There is a very clear development of argument, which is announced at the beginning, and there are logical links between ideas
- Very thorough, detailed analysis with highly convincing, wide-ranging evidence to back up arguments
- Clear evidence of originality in lines of argument, selection of evidence and /or sources
- The assignment is exceptionally well structured and well written, in immaculate, grammatically correct sentence structures and appropriate style, including a full bibliography and references according to the specified referencing style

An answer achieving 80+ must be outstanding in every way. It presents an extremely coherent, well structured, immaculately presented, very well informed argument which is very well supported, exceptionally ambitious in scope, shows originality and relates the topic to the broader context.

Example 3: Extract: There are more categories and also conclusion criteria

Assessment Criteria (Patchwork Text)	Fail 0-30%	Excellent 70% +
Introduction (10%)	Poor introduction that doesn't outline individual expectations of the module and what was considered confusing about it and fails to outline the structure of the report	Excellent introduction that clearly outlines individual expectations and what was considered confusing very clearly. Provides an excellent and extremely clear outline of the report structure.
Main Section A (30%)	Poor commentary that fails to clearly comment on the Task Focus; the Academic Expectations; the MBA content or what was new or unexpected	Excellent commentary that clearly outlines the Task Focus; the Academic Expectations; the MBA content or what was new or unexpected
Main Section B (40%)	No use of Gibbs and no integration of Gibbs as a reference or of the concepts of the Gibbs Cycle of Describe; Examine; Evaluate and Analyse	Excellent and thorough use of Gibbs as a reference and of the concepts of the Gibbs Cycle of Describe; Examine; Evaluate and Analyse

Chapter 14
Can Online Rubrics Develop Learners' Metacognition?
A Qualitative Case Study Analysis

Milena Marinkova
University of Leeds, UK

Joy Robbins
University of Leeds, UK

ABSTRACT

The growing use of rubrics as tools that can enhance students' learning has prompted an accompanying growth of rubric research in higher education, with a wealth of positive findings. As of yet however, these investigations have predominantly focused on paper-based rubrics or their digital static equivalent rather than truly online rubrics, which present a paradigm shift in how rubrics are displayed, accessed, understood, and interlinked with student text and feedback through the digital affordances of hyperlinking. Studies that have investigated online rubrics so far have focused on pragmatic concerns like efficiency or satisfaction with use, which are important aspects of any digital tool, but secondary to learning. The authors therefore carried out longitudinal case studies to investigate what impacts, if any, the online-ness of rubrics had on students' metacognitive development. Results show strong potential for online rubrics to enhance metacognition, but unfortunately in the majority-used platform we investigated, online rubrics currently are more hindrance than help.

INTRODUCTION AND BACKGROUND

While there have been strong criticisms of rubrics for instrumentalising knowledge and limiting learners' cognitive and metacognitive development (e.g., Torrance, 2007; Sadler, 2009, 2014), much of the criticisms have been found to rely heavily on anecdotal evidence (Panadero and Jonsson, 2020). There therefore continues to be a need for careful, empirical exploration of well-implemented rubric use to understand how and where they can improve learning. We are now approaching a level of robustness

in the literature where positive rubric findings of the past 20 years are being confirmed through alternative methodologies. For example, a review of all relevant empirical studies of rubrics for learning (as opposed to e.g., improving marking) conducted in 2013 (Panadero and Jonsson) found noteworthy benefits for students through making assessment criteria transparent thereby reducing anxiety and aiding metacognitive developments such as self-efficacy and self-regulation. However, the same review also found these benefits were hard to credit to rubrics alone given the "instructional interventions", such as teaching of self-assessment skills, that often formed part of the rubric implementation. Nine years later, a tightly controlled empirical study (Krebs, Rothstein and Roelle, 2022) did indeed find that with equal instructional intervention, the use of a rubric as opposed to a non-rubric self-rating schema improved students' self-judgement while at the same time lessening their cognitive load of doing so.

These studies indicate an increasing understanding that rubrics are a useful tool in improving students' metacognitive learning, and yet, despite this positive evolution, there has been no research into how rubrics impact metacognitive development when used in the online space where they, along with much other feedback and assessment, progressively occur. There have been a few investigations into the use of online rubrics, but as is common with research on digital tool use, the affordances of the tool take centre stage rather than consideration of true learning benefit. For example, use of online rubrics has been found to increase teacher efficiency by reducing the amount of time it takes to create feedback (Anglin et al., 2008; Atkinson and Lim, 2013): a worthy benefit, but one we would hope to find from any paper-to-digital switch, and not one that demonstrably improves student learning. Relatedly but perhaps more learning focused, another study (McKinney, 2018) found that a well-planned implementation increased both teacher satisfaction, because by shrinking their marking time it allowed them to give better feedback, and student satisfaction, because the grading was clearer and the feedback more helpful. This starts to make the argument for rubrics in the online space for learning rather than teaching or administrative enhancement, but as any educator knows, increased student satisfaction is not the same thing as increased learning. Another investigation found the assessment analytics possible with online rubrics allowed potentially useful comparisons of feedback and marking practice across a teaching team (Reed, Watmough and Duvall, 2015) which is again a useful benefit (when handled with a sensitive understanding of what analytics mean) but also again one that focuses on teachers and perhaps quality assurance rather than students' development. A likewise tangentially related study that looked at online rubrics used in a MOOC (Ashton and Davies, 2015) found that if students were given rubrics with guidance as opposed to just plain rubrics, they were better able to assess their peers, revealing nothing about online benefits or drawbacks to rubrics (besides the obvious point that online rubrics can be used in online environments at scale) but instead emphasising the previously noted findings that instructional intervention will help rubrics (or indeed any new learning approach) be successful.

The current situation is that despite useful studies into the pragmatic aspects of online rubrics, it has not yet been investigated whether and how the affordances of online rubrics can impact students' metacognition in terms of self-knowledge (i.e., awareness of own strengths and weaknesses, motivations and goals), procedural knowledge (i.e., knowledge of learning strategies), contextual knowledge (i.e., knowledge of academic and cultural norms, tasks and strategies) or self-regulatory processes (i.e., ability to reflect on, monitor and evaluate own learning) (Flavell, 1979; Wenden, 1987; Schraw, 1998; Pintrich, 2002; Rhodes, 2019). This chapter will therefore examine to what extent students' engagement with online assessment rubrics impacts their metacognitive knowledge and control. We will focus on exploring learners' encounters with online rubrics as a series of "metacognitive experiences" (Flavell, 1979) that are multidimensional, dynamic and person-specific (Butler and Winne, 1995), and might

involve varied degrees of affect, cognition and reflection. For the online rubric platform, we will investigate Turnitin Feedback Studio as this is what our students commonly encounter at our institution, a UK university, and is also widely used across the sector (the company's marketing copy claims it is used at 15,000 institutions worldwide (Turnitin, 2021).

METHODOLOGY

We chose case studies as the analytic framework for our research design to fit with the inherently qualitative nature of what we were trying to observe (Creswell and Creswell, 2023) and also as the best way, as viewed in our field, to capture "thick data" on intricate and interacting circumstances (Dornyei, 2007 p.155). Each of the student participants in this research is considered as an individual case, allowing us to follow in detail and observe over time online rubric use "in its natural context and from the perspective of the participants involved" (Gall et al., 2003 p.436). To be sure participants would encounter online rubrics in their studies, we recruited from within our School unit, a Language Centre which teaches English for Academic Purposes (EAP) to students joining our university from non-English speaking backgrounds, as all our courses use online assessment and feedback, including formative and summative rubrics. Students were recruited through promotional talks, and none of our own students were recruited given potential concerns around power dynamics. All volunteers had one-to-one sessions to ensure full understanding of the consent form before signing, and data collection began in the autumn semester of 2019. Given the longitudinal nature of the study and the rising pressures of Covid at the time, there were several dropouts, and we were left with three case studies. We recruited a further two students for a single semester the following academic year (2020/21) to bring the total to five; see *Table 1* for a summary of their basic info.

Table 1. Student participants

Student identifier	Gender	Nationality	Level of study	Future degree
S1	F	Kuwait	Pre-masters	MA Education
S2	F	Kuwait	Pre-masters	MSc Medical Imaging
S3	F	China	Pre-masters	MSc Advanced Computer Science
S4	F	China	Pre-masters	MA TESOL Studies
S5	M	China	Pre-masters	MSc Electrical Engineering and Renewable Energy Systems

Each student participated for at least one semester (3 months) of study, allowing us to observe how they interacted over time with the multiple online rubrics they received. The nature of the intensive EAP courses they were on, with a minimum of 15 taught hours per week, meant they had several feedback points throughout each semester, predominantly on academic writing tasks, some of which included formative feedback on draft work followed by summative feedback on final submissions.

To observe their online rubrics interactions with as minimal intrusion as possible, we asked the students to record a screencast of just their laptop screen (no camera) as they accessed their online feedback

at a time and place amenable to them. In addition, we asked them to follow a think-aloud protocol as they went through whatever online feedback interactions they wished. While this may have impacted the naturalness of the recording somewhat, this impact appeared to diminish over time and also allowed us a necessary window into the students' thoughts and reactions. Students were given one-to-one training in how to make screencasts and carry out think-alouds. To reduce the cognitive load and improve the accuracy of the screencast think-aloud, we encouraged participants to engage in a non-metacognitive verbalisation (Bowles, 2010): they were asked to report what they were doing without necessarily inter-preting all their thoughts, feelings or behaviours while they were doing it (i.e., why they were thinking or doing something, or why they were feeling in a particular way). Moreover, the think-alouds were concurrent in that students were reporting while performing the task (i.e., accessing and reviewing their online feedback), thus ensuring higher accuracy of the think-aloud (Ericsson and Simon, 1993).

No training about rubrics was given as this would have introduced the instructional intervention variable noted in previous studies that muddies the waters in terms of what is causing learning affect. However, some students did receive training about rubrics from their teacher, details of which were as-certained in later interviews. Notably, no students had been given instruction in accessing or navigating the online rubrics. Students were asked to make a screencast with a think-aloud every time they received new online feedback on a task. For 4 of the 5 students, the first feedback point they recorded for us did not use an online rubric – the teaching focus seemed to be more on getting students to submit a short initial task and understand how to use the system to access feedback comments. We purposefully did not specify to students that we were only interested in rubric feedback as that would have influenced the results; instead, we welcomed all screencasts students wished to share with us, and later followed up each screencast, or for each set of screencasts as timing required, with semi-structured interviews to discuss the interactions (or non-interactions) observed on screen.

The interview schedule started with open-ended questions about what the student thought about each task and accompanying feedback, before probing more specifically into what they did with the feedback and what they thought about the rubric. For part of the interview, participants were also asked follow-up questions based on their screencast think-alouds using a stimulated recall protocol (Gass and Mackey, 2016): they were shown stills or footage from their screencasts and asked to retrospectively explain or expand on the thinking behind their online behaviour or verbalisations recorded in the screencast. We wrote up brief contact summary sheets for each student after each initial interview and added to the sheets as interviews progressed to note any salient points and allow for further triangulation when revisiting the data.

The screencast think-alouds (n=18) and interviews (n=12) were transcribed with all data anonymized and each case study's transcriptions arranged chronologically; see Table 2 below for a summary of the data points. The students' written work and teacher feedback were not collected directly but instead viewed via the screencasts, providing additional data as needed for analysis. For this reason, and for other visual and auditory data beyond what was transcribed, the videos were continually revisited dur-ing the data analysis. Although we initially attempted a coding approach to each case study, carried out individually and then comparing between researchers to check interrater reliability, we found that this more formal method created an overly fractured view of both the multimodal nature of the data and of the student-centred narrative enquiry at the heart of the study's purpose. We therefore switched to a broader narrative analysis approach, reiteratively reading each data set to identify each student's themes of feedback use, rubric use, and metacognition. Similar case study analysis is common in English as

Second Language research, and as with, for example, Leki's case studies (1995, 2001), we analyzed not just for recurrent themes but for single instances that stood out due to the students' reactions (2001, p.46)

Table 2. Data gathered for each case study

Student	Semester 1, 2019	Semester 2, 2020	Total
S1	• 3 Screencasts of formative work • 1 Screencast of summative • Interview	• 1 Screencast formative • Interview • 1 Screencast summative • Interview	6 screencasts, 3 interviews
S2	• 2 Screencasts formative • Interview	• 1 Summative screencast • Interview	3 screencasts, 2 interviews
S3	• 2 Screencasts formative • Interview	• 1 Screencast formative • Interview • 1 Screencast summative • Interview	4 screencasts, 3 interviews
	Semester 1, 2020		**Total**
S4	• 2 Screencasts formative • Interview • 1 Screencast summative (of 2 separate tasks) • Interview		3 screencasts, 2 interviews
S5	• 1 Screencast formative • Interview • 1 Screencast summative • Interview		2 screencasts, 2 interviews

It is worth noting that these screencasts and interviews furnish direct quotes in the case studies below, which we have left as recorded or only edited minimally, e.g., to remove repeated words. The reader may therefore notice the quotes do not always follow English conventions they may expect. It is our deliberate choice to use the authentic expression of these multilingual speakers to respect their voices and see what we can learn from them, rather than judging their linguistic proficiency in English as a representative of their whole selves (see Ryan and Viete, 2009 for an analysis of the barriers international students routinely face with this).

THE DIGITAL FEEDBACK LANDSCAPE

Necessarily, the interactions of each student with their online rubrics were mediated by the online platform. It is therefore important to delineate the digital feedback landscape these students navigated. Within the Turnitin platform used, there are three main types of feedback possible, which we refer to in this study as follows:

1. **Overall comment** – a text box which displays down the righthand side of the student work and is always visible, remaining in place as the student scrolls up and down.

2. **In-text comments** – text boxes which are either attached to a teacher-chosen section of text (of any length) or which are free-floating, appearing at a selected point. The student must scroll to the relevant places in the text to see these. They appear as little speech bubbles, which must be clicked to open and read fully, though in some cases a title of the comment type (e.g., Structure) and/or an icon indicating the comment is linked to a rubric may show on the unclicked view if the teacher has used these functions.
3. **Rubric** – the rubric feedback is accessed through a "View rubric" button which is always on display above the Overall comment. (In cases where a rubric has not been used, the button is still present but greyed-out.) Once clicked, the rubric opens in a separate window, usually on top of the student work. It appears as a standard rubric grid with levels of performance across the top and criteria down the side. Where linked Feedback comments have been used, a speech bubble with the number of linked comments displays next to the rubric criterion, and clicking it brings up a list of linked comments which can then be selected to jump to that comment in the work.

In addition to these feedback types, teachers are also able to highlight text, insert text, and cross text out, however these types of feedback were used very rarely in the case studies and are far removed from our focus on rubric interactions: we therefore do not consider them in this study.

As well as these feedback presentations, students are also presented with a grade in the top right if a grade has been given, and a set of clickable icons down the righthand side of their text relating to their "similarity score", i.e., how much of their text is the same as other sources or previously submitted student work according to Turnitin's database. Where text is deemed to be the same, it is automatically highlighted with a different colour given to each Turnitin-identified source. Finally, beneath the similarity icons are two further clickable icons which allow students to view info about their work, and to download it. *Figure 1* below shows the students' main feedback view.

Figure 1. The digital feedback landscape
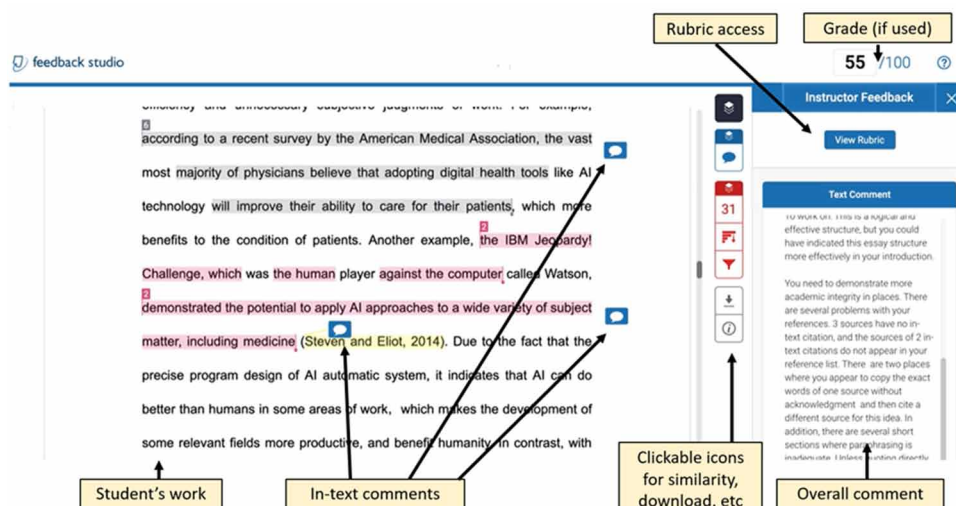
THE ONLINE RUBRICS

Within this feedback landscape, the rubrics were the focus of our investigation, although as noted we did not guide the students to the rubrics or emphasise them, thus allowing students' natural interactions with the online feedback to emerge. All students had access to Word document versions of the rubrics on the Virtual Learning Environment (VLE),though not all made use of this, but within the Turnitin platform where they actually got their feedback the rubric is somewhat hidden, requiring an additional click to open, and once opened it occludes the student work so that the full rubric can be viewed instead. Impacts from these seemingly small digital necessities are discussed in the case studies below. The rubrics themselves were analytic and task specific (Jonsson and Svingby, 2007), though the specificity was at the level of task type, e.g., academic essay writing. They had 5-7 criteria which, in the summative use, each had different weightings that were automatically aggregated into a total score (in the formative, no weightings were present and no score generated). Many of the criteria were "fuzzy", requiring the teacher's qualitative judgement of "abstract mental construct[s] denoted by a linguistic term which has no absolute and unambiguous meaning independent of its context" (Sadler, 1989, p. 124), e.g. "sufficient" vs "good" understanding of the task, though efforts had been made through the criterion level descriptors to sharpen this where possible, leading to relatively lengthy descriptors that noted multiple elements. See *Figure 2* below for an example rubric view from one of the case studies.

Figure 2. Example summative rubric view

Perhaps the most significant online display element for these rubrics was the number of performance level bands or scores. While it is widely acknowledged that fewer – perhaps down to even just three – score levels is most impactful (Suskie, 2017, p.552), these rubrics had many levels, requiring students to scroll right and left across the rubric window to see which cells had been selected. For the formative rubrics, there were 5-6 levels, while the summative ones took all but the highest and the lowest levels and divided them into 3 to allow marking at the high, middle, and low point of the band. This resulted in an unwieldy number of levels, requiring the student to scroll left and right to see them all (note how only 4

of the many bands show in *Figure 2*), but was seen as necessary at the Centre due to external pressures of meeting UK Visas and Immigration language proficiency stipulations and also university numeric grade expectations. Whether these fine-tuned differences could be adequately assessed or represented in the rubrics is beyond the scope of this study, however it is useful to observe students interacting with rubrics that seek to meet the needs of multiple stakeholders, likely not an uncommon experience.

METACOGNITIVE JOURNEY FRAMEWORK

Our approach to analysing participants' metacognitive journey is informed by Flavell's metacognitive framework (1979), in which he distinguishes between metacognitive knowledge of self, task and strategy, on the one hand, and metacognitive experiences as "any conscious cognitive or affective experiences that accompany and pertain to any intellectual enterprise" (p.906), on the other. In this conceptualisation, metacognitive self-knowledge refers to the declarative knowledge of one's abilities and motivations; procedural knowledge refers to general awareness of learning strategies and how to use them (e.g., how to skim or cite); and strategy or contextual knowledge refers to situated understanding of the strategies and knowledge that may be deployed in specific tasks and different contexts (Pintrich, 2010; Schraw, 1998). Metacognitive experiences have a somewhat different function: they act as catalysts, often occurring in situations that "stimulate a lot of careful, highly conscious thinking": they can signal the emergence of metacognitive knowledge into consciousness, or they can add to, revise, or delete from the existing metacognitive base (Flavell, 1979, p.908). Examples of metacognitive experiences might include suddenly feeling clear about a task, unexpectedly recalling having completed a similar task in the past, believing that a problem is easy, or feeling confident about a specific learning strategy (pp.906-8).

Scholars have also drawn attention to the importance of self-regulatory processes in metacognition, which encompass a range of activities that "help students control their learning" (Schraw, 1998, p.114; also, Rhodes, 2019) such as planning a task, monitoring task performance, revising goals, or evaluating the effectiveness of learning strategies. In their model for self-regulated learning, Butler and Winne (1995, see *Figure 3*) identify the importance of feedback – both externally generated by tutors and peers, and internally generated by learners themselves:

Figure 3. Butler and Winne's model of self-regulated learning
(1995, p. 248)
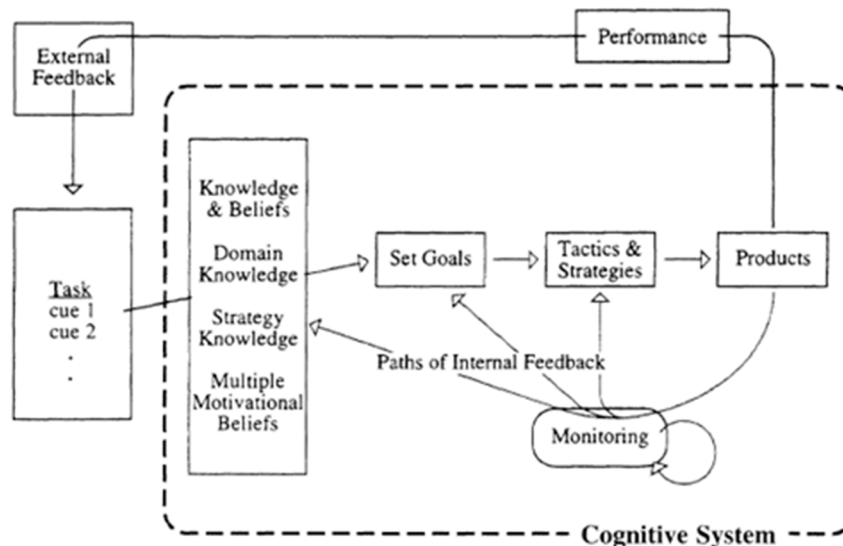
They argue that the quality of external feedback, and this would include online rubrics, can impact learners' affective dispositions towards their perceived achievement, which in turn might influence – positively or negatively – their monitoring of goals (self-knowledge), adjustment of strategies (procedural knowledge), and specific task engagement (contextual knowledge). For effective self-regulation to occur, therefore, feedback practices should be developmental and dialogic, not just outcome-oriented, in order to help close the gap between current and desired performance, provide clear guidance to ensure uptake, and take into account the role of learners' prior knowledge and beliefs about the subject, learning and themselves as these might skew how formative messages are interpreted (Nicol and Macfarlane-Dick, 2006).

What is often highlighted in conceptualisations of metacognition and self-regulation, as well as their relationship to feedback, is the "recursive" nature of the information flow and knowledge construction processes:

students may modify their engagement by setting new goals or adjusting extant ones; they may reexamine tactics and strategies and select more productive approaches, adapt available skills, and sometimes even generate new procedures. If external feedback is provided, that additional information may confirm, add to, or conflict with the learner's interpretations of the task and the path of learning. As a result of monitoring task engagement, students may alter knowledge and beliefs, which, in turn, might influence subsequent self-regulation (Butler and Winne, 1995, p.248).

Given the multiple and complex processes involved in self-regulation, Butler and Winne (1995) acknowledge that existing empirical research on feedback and regulatory processes tends to aggregate data and thus potentially "fail to reflect the variance in [learner] behavior" (p.246) or the dynamic nature of metacognitive knowledge and self-regulation. Therefore, our approach to the collected data on students'

use of online rubrics was exploratory, i.e., we were interested in mapping out our participants' multiple encounters with online rubrics as metacognitive experiences that might have triggered or inhibited the emergence of a particular type of metacognitive knowledge and self-regulatory behaviour. At the same time, we also aimed to acknowledge the differential impact online rubrics had on our participants' metacognition and the dynamic nature of metacognition and learning in general. Thus, we traced each participant's journey over time, which might have not always been linear or progressive (as Butler and Winne have argued). Our analysis of each case study therefore traces the metacognitive journey by starting with the initial state of their self-, procedural and contextual knowledge and self-regulatory processes, moving on to describe their metacognitive experiences of online feedback and rubric (non)use, exploring the rubrics' impact on the learners' metacognitive knowledge development, and identifying any adjustments to, revisions of, or persistence with metacognitive knowledge and behaviours.

The richness of the data combined with our non-directive, enquiring approach allowed us to trace multiple metacognitive experiences in each student's development. Inevitably for such a large, qualitative data set, many of these metacognitive experiences were not directly rubric-linked, but still formed a necessary part of the larger case study picture that allowed us to build rapport with the student and also let their online rubric use emerge naturalistically, a key goal given previous shortcomings noted in the literature. Each case study's rubric-related metacognitive experiences are given below, in the order presented in *Table 1*.

S1 CASE STUDY: COMPETENT RUBRIC USE BUT METACOGNITIVE GROWTH OFFLINE

S1 was a confident, aspirational student with previous teaching and publishing experience, and the ambition to apply for a doctoral study after her Masters'. She had already been studying EAP at the Centre for one semester previous to this research, and was committed to the project, recording additional screencasts and volunteering for project activities. S1 was aware of online rubrics from the start of her involvement with the project, being the only participant to refer to them in her screencast think-aloud upon receipt of her first rubric-referenced feedback. During the project, she developed this awareness of rubrics into a deeper understanding: from seeing rubrics as an important codification of teacher expectations and achievement levels to a tool for prioritising learning strategies and monitoring of progress in different academic literacies. This metacognitive growth, however, was not necessarily down to online rubrics; as S1 was to discover, their inaccessibility pre-submission and limited functionality post-submission often made it impossible to engage in the type of learning, self-knowledge, and self-regulation she was after.

From the start, S1 demonstrated assured declarative knowledge about herself as a learner, and unlike other project participants, and a significant proportion of EAP and university students more generally, S1 read into the tutor feedback a fairly balanced reflection of her abilities rather than seeing it through an exclusively deficit lens (Benesch, 2001). In her first screencast think-aloud, for instance, S1 showed insight into her learner self by agreeing with her tutor's feedback that grammar was her weakness, but also challenging her tutor's assessment of her reading skills as she knew this to be a strength, just one that her writing did not yet show:

I have some issues here and really I know about them. Structure of sentence and this is my weakness. …
[Reads tutor's overall comment.] "You also do not seem 100% clear on what the article is saying." Uh,

really, I do my best, and I think I understand the reading of the articles. But the difficult when I want to express my point.

S1's self-knowledge was also punctuated by feelings of intellectual challenge, and investment to improve. This range of metacognitive experiences was further supplemented by S1's confidence, which she showed by not merely reading verbatim her tutor's comments but by responding to the feedback, a sign perhaps of her pre-existing high self-esteem, recognition of the benefits of praise, and willingness to learn from past mistakes (Young, 2000):

[Reads tutor's overall comment.] "Good. You appear to have read a lot, and have tried to synthesise these sources...." Actually, really it is difficult to organize the idea because the thematic essay is need to simplify the idea and organize the paragraph in logic. So, I think here I have to improve... "You need to simplify your organization. This may mean you rewrite an outline." So, what I have to do to recognize all these details and try to improve and rewrite the outline. All this. Because, really, I have good sources… But I have to just re-outline.

What can also be seen is that S1 did not merely report experiences such as challenge, effort, or confusion that the encounter with feedback had triggered. Rather, she drew on these as a springboard for deeper learning and motivation, which has been linked to improved self-regulation (Zimmerman, 1990). In other words, her initial confidence about her successful use of sources seems to have provided S1 with a clear sense of direction on what to work on next (structure), and possible learning strategies (rewriting and outlining).

It was at this point that S1's first experience of online rubrics occurred, even though it was limited to identifying her performance level for each criterion:

And here we have "View rubric". It's important to know where's my position. So, in the… first draft, I have "Some way from Pass" on the whole text. OK. And I have also for the "Development of idea/ argument" … I need to work hard just to achieve… the "Pass" level.

While identifying her achievement level ("Some way to Pass"), she also begins to use the rubric for self-monitoring, noticing her "position" against individual criteria. This reviewing of criterion scores has been noted as a beneficial approach to rubric use for self-regulation (e.g., Andrade and Du, 2005). Furthermore, the encounter with the online rubric seems to have prompted S1 to begin self-regulatory reflection on the amount of effort required to improve in these specific criteria ("I need to work hard just to achieve ...the 'Pass'"), in a way not dissimilar to her earlier comments about her drive to do well, to work hard and address challenges, which revealed her growing self- and procedural knowledge.

This somewhat basic online rubric use post-submission was complemented by S1 using the rubric pre-submission to inform her work, something she had not done in the previous semester (before this study). As she explained in the follow-up interview, S1 had been made aware of rubrics – tutors "show us a file for all documents we need for the essay, and they explained everything" – but she did not make use of them at that time as she "didn't know how to work with them", confirming other findings in the literature that instructional intervention is often insufficient to promote informed rubric use by learners (Brookhart and Chen, 2015; Johnsson and Svingby, 2007; Panadero and Johnsson, 2013). At this point in her studies, however, S1 was able to engage, downloading the Word document version of the rubric from

the VLE and using it from the start of her writing process to "focus on each part [criterion] to reach all the goals they [tutors] need". Metacognitively, this shows her evolving procedural knowledge for rubric use, but importantly for our focus, not one that uses online rubrics. Due to the lack of functionality of the feedback platform to enable the pre-submission use of interactive rubrics by learners (Robbins and Marinkova, in press), this element of S1's rubric-driven growth happened all in the static, downloaded document space.

When S1 did make use of the online rubric, she bypassed its interactive features despite her ease navigating the rubric space – using the scroll bar to see the selected score levels, moving the rubric around, zooming in and out of the rubric. In her screencast think-aloud (see *Figure 4*), it can be seen that the in-text comments in S1's script were hyperlinked to specific rubric criteria, which should have enabled positive "synergy effects" of the two feedback channels: the overall achievement level and key areas for long-term improvement explicated via rubric criteria should have worked in tandem with the specific instances that require immediate action highlighted in the in-text comments (Nordrum *et al*., 2013, p.930). However, when S1's cursor hovered over the online rubric criteria and the hyperlinked in-text comments appeared briefly underneath, she did not take any notice. In this sense, the potential of online rubrics to provide the in-task synthesis of different feedback elements was not realised. Instead, S1 appears to have approached the online rubric as a separate digital space: she proceeded to read out loud each highlighted performance level without linking the criteria to tutor in-text comments on the script (or overall tutor comment, for that matter). This separation was compounded by the need for the online rubric to open in a separate window, which covers the student script and other feedback – the hyperlinking functionality failing to overcome this separation.

Figure 4. S1's view of in-text comments hyperlinked to online rubric criteria

For S1, then, the onlineness of rubrics did not seem to have played a major part in her metacognitive journey – or at least not in a positive way. When reflecting on her use of online rubrics pre- and post-submission, she pointed out that the static version had been more conducive to her learning and regulation before submitting a task: "I prefer to download and print the copy and determine to read more in details and make highlights for some section that I have to focus." When S1 downloaded the online rubric post-submission to monitor her feedback, it did not appear as a table but as a very long list of descriptors, which made it hard to make sense of the academic expectations at different performance levels. The student also struggled taking screenshots of the rubric to be able to use it simultaneously with

her script – the view of the latter being blocked when the rubric opens in a new window – thus preventing the use of the rubric to identify specific instances of strengths or challenges, or to monitor recurrent errors. In fact, it was only when S1 worked with the rubric offline or in hard copy that she was able to analyse criteria and apply them to her work pre-submission for self-assessment, or to her achievement level post-submission for progress monitoring:

I prefer to save all document in one file just to see the improvement… Folder for Level 2, and folder for Level 3… for all feedback together… I download the PDF. Feedback overall. And …the comment in details… I put all their recommendation in one file. And… make a schedule to plan how to improve myself and I list the main problem because there are some repetition of the mistake... and there is some main important skills in writing that I have to improve also. Even if I'm good in it, but I have to be focused and not forget… how to write.

It was the folder she had created in the offline / paper space that enabled S1 to integrate all elements of her feedback – overall comment ("Feedback overall", "recommendations"), in-text comments ("comment in details") and rubric – to monitor her progress not only within a task ("see the improvement"), but also across assignments and courses ("Folder for Level 2, and folder for Level 3"). In contrast, and somewhat paradoxically, in her encounters with the interactive rubric, S1 would approach it as a separate space; to integrate it with other feedback elements, she had to leave the online medium.

For all the limitations of the online medium, rubrics in general seem to have had impact on S1's metacognitive growth. The participant regarded her initial approach to rubrics as limited – she "downloaded [the rubric file] at the beginning of this semester, just to understand what they [tutors] need". However, by the second semester of this study (which was confusingly called Level 4 for her), in addition to using rubrics to develop her assessment literacy pre-submission, and to monitor her progress post-submission, S1 seems to have noticed their potential to be a useful "instruction" to learning:

Actually, in Level 3 I didn't consider importantly for the rubric. In Level 3, I just focussed … "Yes, this is score, it will be fine". But I didn't at the beginning to consider the instruction. But really, I found this is important to help me.

In terms of their motivational impact on self-knowledge, rubrics seem to have led to a shift in S1's metacognitive experiences. If in earlier encounters, her online rubric-enabled self-positioning brought a degree of satisfaction at achieving a grade ("the score"), at the end of the semester S1 voiced an awareness of rubrics' learning potential, but a potential which is learner- rather than tutor- driven: with students looking back at earlier assignments, comparing achievement and feedback, devising appropriate learning strategies:

At the end, I saw how is the progress of the feedback. And when I returned back for the first draft, and I compared it with the final, we can see how they improve the mistake and academic writing. And this one it depends on the student, how they are dealing with the feedback.

In this sense, therefore, it was S1's response to the often unwieldy online rubric, i.e., move offline, download, collate all feedback, which showed the student's metacognitive knowledge in all dimensions – of self, of context and of self-regulation. S1's self-knowledge was enhanced as her offline approach

to rubrics helped her notice patterns in tutor comments, identify areas of strengths and challenge, and devise effective learning strategies. It might be argued that the layout and accessibility limitations of the online rubrics pushed S1 into developing further her procedural and contextual knowledge of how to usefully deploy these particular online rubrics. And finally, S1's decision to collate her feedback from various assignments to monitor her performance demonstrated her self-regulation, an approach over which she had more ownership than the feedback platform would allow. These conclusions echo recent challenges of the allegedly "representative" function of rubrics, seen as offering absolute transparency of what is otherwise opaque "institutional intentions" and processes (Bearman and Ajjawi, 2021, p.362), a promise reiterated even louder by the interactive affordances of online technologies. Instead, after having encountered a number of barriers to the promise of transparency and failing to notice the interactive possibilities, S1 decided to create her own "productive space" (Bearman and Ajjawi, 2021, p.363), albeit using static / paper technologies, where she was able to make sense of criteria, use feedback to map progress, and advance her learning.

S2 CASE STUDY: RELUCTANT AND INSTRUMENTAL RUBRIC USE

S2 was an engaging, funny student who expressed herself strongly throughout the study: there was much she did not like about the online feedback she received, and she was very comfortable sharing that in her screencasts and interviews. She was extrinsically motivated, noting her wish to get a high grade, but she also took feedback very personally (negative feedback made "all the weekend bad"), revealing a deep care for her work and a wish for her efforts to be understood and acknowledged, ideally praised ("I do hard work, I need someone to appreciate me not to told me... you didn't understand"). Her struggle with the affective dimension of feedback literacy, her ability to "harness the emotional impact of feedback" (Winstone and Carless, 2022, p. 27), was one of the main barriers to her developing higher metacognitive skills of self-regulation. In the first semester, this barrier took the form of negative reactions to the negative feedback she received, while in the second semester, with a new teacher who had a more positive feedback style, it took the form of a continued reliance on the teacher to tell her what to do. Although she was aware of the existence of the assessment rubric from the previous term (before this study commenced), she did not use it at all in Semester 1 because her teacher "didn't send for me". This shows her teacher-responsible view of learning, as the students were expected to access the rubrics through the VLE and through the feedback platform, but also indicates a possible barrier in the online rubrics' lack of visibility. Rubric awareness only emerged in the second semester, and never formed part of her learning or metacognitive development.

Early on in semester 1, S2 went through a significant and relatively common metacognitive experience in many students' lives: she received feedback indicating her work was not as good as she had thought it was (see Rhodes, 2019). In her first screencast, on a casual writing task that she understood to be low-stakes, this disjuncture between expectation and experience did not occur; there were both positive comments and relatively minor suggestions which she was able to take in stride and respond well to: "I like this one feedback, something to encourage me. ...Thank you. ...Okay. Work, I will work." While these responses show undeveloped learning strategies ("I will work") and a very basic emerging self-knowledge about weaknesses, they nonetheless evidence S2's ability to reflect on non-threatening task feedback. However, by her second screencast, with a weightier academic writing task at hand and a preponderance of negative feedback comments, she responded instead with an argumentative and

unhappy tone: rather than engaging fully with the Overall comment – which is also the place where she could have accessed the rubric for further feedback on strengths and weakness of the overall work, she quickly skipped over it and instead moved to the more granular In-text comments. In the screencast, she did not access the rubric at all and instead worked through the in-text comments one by one, often challenging them as wrong and then moving on quickly:

What he wrote for me? [Clicks on in-text comment, reads:] "The form of the verb used is incorrect." Why? This sentence exactly I did before and the same teacher told me it's OK. Why it now wrong?

[Clicks on next comment, reads:] "There is missing words." Not clear for me.

As per Butler and Winne's (1995) model of self-regulated learning, the perceived quality of the external feedback had a negative impact on S2's affective disposition, creating an initially negative metacognitive experience (Flavell, 1979). Despite this, the follow-up interview revealed that after the initial feedback interactions she recorded in the screencast, S2 then returned to the feedback multiple times at different points, printed out her work and annotated it based on the feedback, and made an appointment with her tutor to discuss. The full metacognitive experience therefore shows both that she is beginning to develop the contextual knowledge of how to navigate academic culture and strategies for academic writing, and that the challenge to her self-perception of her work has begun to evoke nascent self-regulatory processes and self-knowledge. In the face of such a challenge, it is unsurprising that S2 did not have the mental energy to take in other elements of the digital landscape, in particular the View Rubric button. The online-ness of the rubric in this case made it completely occluded: either she did not see or understand how to use the button (most likely given her unawareness of the rubric revealed later in the interviews and also her low digital proficiency displayed during the initial one-to-one sessions for this study), or she chose not to click on it because she did not value the rubric.

In the second semester her screencast think-aloud had a much happier tone, with explicit gratefulness to her new teacher who gave feedback that she confirmed in her interview was much more her preferred style (very positive, explicit about corrections). Throughout the recording S2 responded positively to this difference, e.g., "Thank you. I'm happy now. This teacher is different [than] before. Always support me and improve my spirits. She always optimistic." While the affective dimension of her feedback experience was therefore much improved, she did not appear to show metacognitive development in terms of self-knowledge or self-regulatory processes, being dependent on this positive kindness to point out her strengths and weaknesses and to directly tell her how to fix errors, nor was there any demonstrable metacognitive growth in her procedural knowledge, such as learning strategies, as her go-to strategy when she did not understand the feedback remained to ask the teacher:

sometimes [in-text comments] like, "punctuation", or something "word" or "verb", I'm confused I don't know where's the answer. So, I'm waiting for the next meeting to discuss this point with my teacher.

In this lack of development, she confirms Hattie and Timperley's (2007) combined meta-analyses findings that personal feedback or praise, though much liked by students, is often an ineffective type of feedback, rarely creating learning gains. Indeed, this more positive second semester feedback reaction cannot truly be characterized as a metacognitive experience as it did not seem to "stimulate a lot of careful, highly conscious thinking" (Flavell, 1979, p.908), but rather good feelings and continuation of current practice.

Nonetheless, in this semester she did engage with the online rubric for the first time, accessing it in her screencast after the overall comment, in-text comments, and grade to determine her performance level for each criterion:

OK, a "Distinction". [Points at the "Argument" criterion, reads some of the description out loud]. Thank you, thank you for this one grade; I think it's fair for me. …Oh, I have something here in "Distinction Plus". [Points at "Grammar accuracy" criterion]. I'm happy to get this one final result. OK, I'm between "Distinction" and "Plus Distinction".

In the follow-up interview, S2 stated that the rubric "is not more useful than the comment. But it is OK", showing minimal surface level engagement only, which she further refines by saying "if 'Distinction', I like it", revealing a wholly extrinsic motivation (Deci and Ryan, 1985) to using the rubric. However, such extrinsic focus is often present on these EAP courses, where students must achieve a certain level to pass onto their future degree, a context which S2 is well aware of: "it's the course, you know, the mark, it's important thing to us on the course." She elaborated on this motivation for her personally later on, noting the rubric's purpose in this use: "I read to see I'm 'Distinction' or I'm 'Merit', like this, because I have to pass not [the] same [as] people, I have to be more than another [in] this level." Whether this motivation is "external regulation" due to her need to attain a higher grade than her peers going on to programmes with lower entry requirements, or "introjection" based on an ego-driven desire to outperform her peers (Ryan and Deci, 2000), it still brings her rubric use into an understandable but unideal instrumental approach.

Specific questioning by the interviewer confirmed that the online rubric had not formed part of her writing practice (and she did not engage with the static, downloadable version on the VLE at all). Her use of the online rubric was entirely instrumental: she only looked at it post-submission to see the mark. Her screencast showed as well that, much like S1, she accessed the other two feedback elements first, the overall and in-text comments, and opened the rubric last for a one-off viewing.

While the low visibility and separateness of the online rubric may have been part of the problem, the real issue was rubric design: the lengthy descriptors and the unwieldy repetition of the same descriptors three times for the main performance levels (see Online rubrics section in Methodology above) was clearly a driver in her highly instrumental rubric use. As she explained in her final interview:

to be honest, I don't read this one sentence [referring to the rubric descriptors]. Just I saw where is the line for "Distinction", "Merit", like this. …I'm not concentration about what's in this [rubric cell] because so many sentence. And I've told you before, it's not [more] useful than the comments and the correction code. …So many sentence and boxes for slides, same levels or same "Distinction" or same "Merit", like this. So many.

Though the confusingly numerous, repeated descriptors in the summative rubric were a result of it being online in this particular platform, it is possible to design simpler rubrics in the online space, and so this barrier cannot be said to be a fault of online alone. S2's confusion can therefore be attributed partly to poor rubric design and partly to platform-specific problematic design aspects.

In sum, for S2, the online rubrics were not a success. Their occluded visibility in the online space seemed to pose a barrier to use for a significant portion of her studies, and once she did start to use them, their online-influenced poor design meant she did so instrumentally; they did not promote any

metacognitive development. To be fair, S2 did not engage with rubrics when explained in class or made available for static or offline use in the VLE either: rubrics overall had very little presence in her learning journey. Yet it is shame that the multiple "invitations" to learning that rubrics can promote (Bearman and Ajjawi, 2021) were essentially quashed by various online barriers in this case.

S3 CASE STUDY: DEVELOPING RUBRIC USE BUT METACOGNITIVE GROWTH HINDERED BY ONLINE

S3 presented as a very polite, dedicated student, perhaps a little shy, who developed well over the two semesters, with her think-alouds, interviews, and onscreen work all showing increasing awareness of academic literacies that take a long time to bed in, e.g., structuring longer work. She was worried that her language ability would be too low for us to understand in her think-aloud recordings, revealing a common discomfort that international students face when dealing with the "nativespeakerdom" of the English-speaking academy (Ryan and Viete, 2009). Despite reassurances, she chose to record her screencasts a day or so after first accessing and processing her feedback, rather than concurrently, so that she could speak more ably, leading perhaps to more reflective and less emotional recordings than might otherwise have been, but nonetheless useful windows into her feedback engagement process and furthermore ones that worked for her individual needs and preferences. Also of note in her screencasts is her adeptness at navigating the digital feedback landscape: unlike S1 and S2, she used the similarity tools via the sidebar icons (see *Figure* 1), going on to describe them with understanding despite, as confirmed in her interview, not being taught about them and simply deciding to click and explore. However, even with this technological proficiency, she did not click and explore the rubric in her first screencast, only finding it when a friend pointed it out to her, and, in common with the previous two case studies, preferred to use paper for working with her feedback:

I think paper is so convenient. I can take some notes in the paper. And I can write anything I think. I think it's so convenient but, in the computer, I think I just look at, I don't take notes. …to write some articles on the computer is okay. But when I analyze something to take notes, I want, I prefer to print.

It is notable that, as with S1 and S2, the feedback platform was for S3 an overly one-way feedback delivery system that pushed her into an offline space where she could actually engage. Similarly, though her awareness of the rubric developed greatly over the course of the two semesters observed, with a clear accompanying metacognitive growth, this appeared to be linked to rubrics alone rather than their onlineness; in fact, the online aspect was frequently a barrier to the rubric interactions S3 wanted to undertake.

S3's emergent rubric understanding was part of her metacognitive journey as a recipient and user of feedback, and it is important to note that to this she brought a significant strength which was unique among the case studies: she had already learned in her previous studies, and perhaps had a natural affinity for, seeking out peer feedback. As she noted in her very first interview when discussing actions she took from the feedback: "I also to talk about my question, my problems with my friend. And he also received the feedback, and we compared [each] other's feedback and we will also talk about how to write this article better."

Indeed, it was only through her friend's prompting that she engaged with the rubric at all. Her broader feedback journey took a traditionally bumpy-but-positive path, largely helped by her self-regulatory

processes and ability to manage the affective domain of feedback. While first S3 found the feedback "too tiring" and it made her "feel sad", she was able to recognize the cause and respond well: "before… I think my article is great… but when I see this feedback, oh, yeah, maybe I'm not so good… I have some problem I should solve." This well-noted metacognitive experience of thinking one's performance is stronger than it actually is (Rhodes, 2019) can be debilitating for students, as it was for S2, but S3's problem-solving emphasis here indicates an ability to adjust self-knowledge of weaknesses and manage the negative emotions of doing so. By the second feedback point, she approached the feedback with trepidation ("First, when I see the so long comments, I am afraid like my essay have a lot problem") but after reading, she took on a much more productive and positive stance, noting that "Yes, I have a little problem, but… have some solution to address this. It can be solved, it doesn't matter." By the third interview, she was carrying out a sophisticated self-regulatory process of monitoring her learning and development by comparing feedback from previous drafts to feedback on her final submission. Admittedly this development was scaffolded by the built-in feedback cycle of course, with feedback first on a draft followed by on a final submission, but not all students take on this learning point even when designed into the curriculum. By the end, she seemed to have reached a new level of metacognition touching not only on her self-regulation, but also on contextual understanding of the academic culture around writing expectations for postgraduates in the UK:

from the feedback, I know the importance in that the tutor is just to… guide you how to study; is just give your suggestions, is the most important thing always to do by yourself. You don't… rely to other people to help you how to write; they just to give your suggestions.

This impressive development of feedback literacy included a strong development of her use and understanding of rubrics. At first, she evidenced very little awareness of them. She did not access the rubric in her first screencast but did mention it in the follow-up interview after learning about it from her friend, noting "He taught me, but I don't attention to this, I don't know. He told me… what's the score you get, what's the band you have?" Understandably with this introduction, S3 saw the rubric as merely a tool for goal setting, stating that "I want to lift up to 'Merit' or 'Distinction'."

After mentioning the rubric at the start of her second think-aloud, she did not actually open it, instead spending 14 ½ minutes (longer than her other screencasts, and the longest of all the case studies) to carefully explain first an overview of her tutor's overall comment, then work through the in-text comments and also note the Turnitin similarity text highlights, before returning for a more in-depth engagement with the Overall comment. Her calm, methodical approach showed clearly that she valued the comment feedback, taking time with all of it, and also just as clearly that she did not yet value the rubric as it got the barest of mentions in this highly detailed account. In the follow-up interview, a question about the rubric evoked almost an apology: "sometimes I, I really forgot to click the rubric. But yesterday my friend told me… yeah, so I click it. But sometime… I really forgot to see it." When pressed for specific pros and cons of the rubric, she responded positively to the clarity, noting that it was "clear to read and so detailed. And different parts have all their own different criteria, have different standards", which appears to echo the finding from Krebs *et al*. (2022) that rubric-arranged feedback can lessen the cognitive load of parsing feedback. If anything, S3 wanted sharper divisions in the rubric, asking for specific scores for each band rather than just an indicated range so that the scores could "remind you what level you are in and what problem for you." This overly marks-focused suggestion, particularly for formative

work, shows that she still viewed rubrics as an assessment rather than a learning tool, but at least with potential for formative use.

In her third and final screencast, on feedback to a lengthy summative essay which previous work fed into, she did open the rubric and talk through it, yet the online display of the rubric, in this instance showing only 4 of the 7 criteria on screen and requiring scroll to see the rest, presented a barrier: S3 never accessed the final criterion in this screencast, despite the her highly-detailed, explanatory approach. Nonetheless, her think-aloud shows the rubric prompted useful self-knowledge development as she engaged with her performance levels for each of the different criteria:

for the "Conventions" I got 48 … is "Fail". I know, my reference, I have some errors… citation is not very well. So, from this I know what's wrong.

And for the "Sources" it's just a "Pass". I think is not good. So, I will have a long way to go in this part.

And the "Cohesion" and "Grammar" is higher than other parts. I'm very happy and I, I didn't think I have a higher score in this part, ...because I always think my grammar is poor, is not good. So, in this essay… I have a higher score in this part. I'm very happy. And thanks my tutor... give me my confidence for my grammar in the future, but I will improve it.

These useful rubric-prompted reflections are not an element of the online aspect, however.

In the final interview, S3 indicated that the rubric had become a tool for self-guidance, explaining how she used it after writing the draft to compare the criteria to what she had written and make changes accordingly. Here she displayed deep, rubric-influenced metacognitive growth similar to her overall feedback literacy development outlined above: firstly, the rubric positively influenced her self-knowledge as she used it to help identify areas of weakness to improve, and to motivate her rewriting processes; secondly, the rubric led to advances in her procedural knowledge of strategies to improve her academic writing as she had to explore various approaches to improve on the criterion she found most difficult, "Grammar". Notably in all these developments, the rubric she used was not the rubric in the online feedback platform as this was viewable only within the platform on submitted work, and additionally printing it out in attempts to bring it into her preferred learning environment produced a multi-page list rather than the onscreen tabular format expected. Just as with S1, S3 understandably found the incredibly lengthy printed list "not easy" to deal with, going so far as to politely suggest in the interview, "if it's okay, I want the rubric can be print." Furthermore, she was unable to use her established approach of seeking out peer feedback when it came to the previously received rubric feedback as the platform barriers prevented it: she noted she wanted functionality wherein "the rubric can add or share links with others" but that this wasn't possible. These online barriers meant that the rubric she used was the accessible, static Word version from the VLE which she could easily download for use on her computer and print for use offline.

The barriers to online rubric use did not prevent her metacognitive growth, however, and by the end she showed an excellent metaphorical understanding of the difference between the personable, directly focused feedback in the tutor comments versus the more impersonal, broader feedback of the rubric:

I think the comment is just some my tutor's views for my essay. I think his opinion is to get me to write a better essay and the rubric is more detail to teach you how to do this. The tutor is just a guide, guiding; the rubric is the way.

This metaphor seems to us an eloquent proof of the metacognitive growth possible through rubric use, yet clearly in this case this growth was due to every element of rubrics except their onlineness.

S4 CASE STUDY: LIMITED RUBRIC USE WITH METACOGNITIVE GROWTH IN A DIFFERENT ONLINE SPACE

S4 was a lively and enthusiastic student, with a well-developed feedback literacy from the start. She appreciated feedback and recognised its developmental impact beyond assessment and grading: "The grade is not everything. What's important is about the detail; it's the feedback itself". This student was also eager to enhance her contextual understanding of different feedback practices, which was reflected in her recording additional think-aloud screencasts, beyond the basic requests we made for the study. More online and procedurally oriented than S1, S4 created her own Personal Learning Environment (Haworth, 2016) in OneNote, where she engaged with feedback in dynamic, flexible and student-centred fashion, and which helped address her specific learning needs. Despite her advanced feedback literacy and metacognitive growth, S4's use of online rubrics was limited, even though she was well-aware of them and was keen to make the most of their learning potential. Thus, her non-use of the institutional online rubrics and preference for an alternative feedback space were perhaps indicative of not only "what was not there" but also of "what could have been".

From her first think-aloud, S4 demonstrated good metacognitive understanding of the factors that can positively impact her motivation and learning – in her case the affective dimensions of feedback (see Hyland, 1998; Pekrun *et al.,* 2002; Pitt and Norton, 2017; Yang and Carless, 2013): "[*Tutor's name*] is really good at encouraging us … whenever I saw all this kind of praise…, I really feel proud of myself and to establish my confidence in writing or in learning English".

Her awareness of what feedback and learning strategies might work for her was often situated within the wider context of the feedback practices at this educational institution. While appreciative of her tutor's sense of responsibility and commitment to teaching, the student also showed awareness that this approach might be qualitatively different from the one during her future MA studies, which in turn increased her motivation to make the most of any feedback opportunities:

My tutor… takes her responsibility to fulfil the feedback. It's very, very detailed. I really appreciate it. So, it gives me a general knowledge of what I should keep, some of my merits, and what should I avoid, some of the mistakes… Maybe in the future… in my master's course, my tutor can't check so carefully... grammatical mistakes or vocabulary mistakes. Maybe he or she pays more attention on the topic or the direction of my dissertation. So, before that time, I need to consolidate my essential skills in writing assessments.

Importantly then, developing the right learning strategies was something that S4 sought from very early on. In response to her tutor's initial feedback on a short task and the challenges identified therein, e.g., academic language, paraphrasing and referencing, S4 adopted strategies that were deliberate and dynamic compared to the somewhat more generic and non-transferable procedural knowledge demon-

strated by other participants in the early stages of their metacognitive journey. S4 created a checklist of her strengths and weaknesses in OneNote, where she would record tutor feedback, her own examples of relevant language and discourse features, as well as personal reflections on the tutor comments and her learning (see *Figure 5*):

Figure 5. S4's feedback portfolio in OneNote
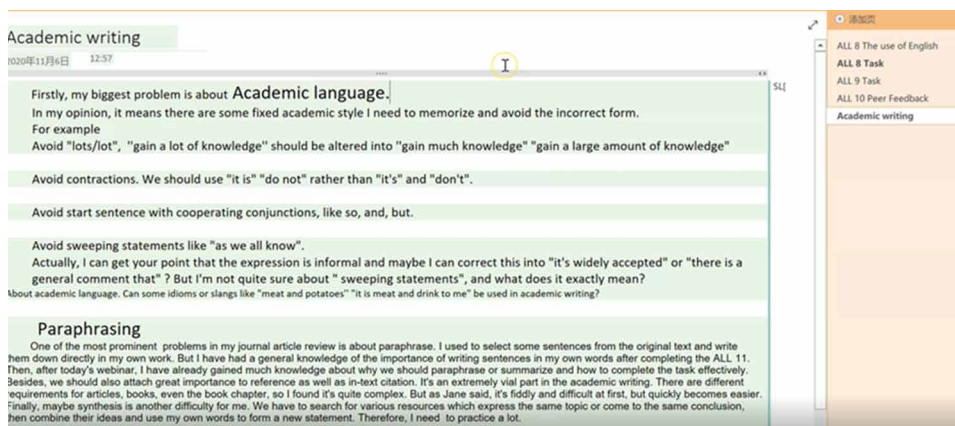
In her screencast think-aloud, S4 elaborated on her choice of OneNote, which has facilitated her meta-cognitive growth in all dimensions – of self, context and self-regulation. It enabled her to organise the feedback in a way that made sense to her learning, to add her own reflections, and thus actively process tutor advice, to insert URLs to external resources providing insight into S4's new academic context:

So, when I research... all the problems in my article, I make a checklist here. This is my OneNote. … This tool is really useful and appropriate for me 'coz it can meet all my demands in learning... Like this: Week 1, Week 2, and in Week 3… I summarised some of my problems here and in order to avoid the same mistakes next time. So firstly, it's about my academic language. It's the first time I exposure the academic style, so it's really useful. For example, I need to avoid "lots" or "lot", so it should be altered into "much" or "a large amount of" … So, after this, I summarised some tips. For example, I can understand the original work deeply and thoroughly, and then I can write down this opinions or ideas by my own words. And lastly, about my references here; it has so many requirements for different kinds of, you know, like books, book chapters and even articles. So, [tutor's name] provide this website to me and I found this extremely useful.

In addition to these affordances of a OneNote portfolio, in her follow-up interview S4 commented on the platform's metacognitive benefits for noticing patterns in errors: "it's organised by myself... the same mistakes appeared many times in my one assessment, so I can conclude them together and give each problem a subtitle … and then I can make such a checklist". Even though at this point S4 did not refer to the institutional online feedback platform, which is much more tutor-led and monologic (Robbins and Marinkova, in press), her choice of an alternative suggests that she, similar to S1's move to a paper-based folder enabling cross-task synthesis of feedback, was after an online "productive space"

(Bearman and Ajjawi, 2021) that could meet her specific learning needs and allow for her own active understanding of feedback practices to emerge.

S4's attempt to navigate the limitations of her online academic context to feedback as a student-led and dialogic practice (Bloxham and Campbell, 2010; Nicol and Macfarlane-Dick, 2006; Yang and Carless, 2013), was also evident in her pro-active stance seeking feedback. When submitting her formative assignment, she initiated a discussion with her tutor and added in her Word document a request for more targeted feedback (see *Figure6*), to which the tutor replied by using the in-text comment facility. S4's request in this instance both made use of **and** challenged the rather monologic functionalities of the institutional feedback platform, which are intended for delivering an old-paradigm type of feedback, or "feedback as information" (Winstone and Carless, 2022). Instead, S4's request ensured that the feedback she received made sense to her and was therefore accessible for uptake and enabled the adjustment of her learning strategies (Carless and Boud, 2018). S4's decision therefore not only shows her advanced metacognitive understanding of feedback as a dialogic process, but also her contextual knowledge of how feedback practices operate in this academic setting and her desire to shape the institutional feedback platform into a space that "can meet all [her] demands in learning".

Figure 6. S4's dialogic request for feedback in the institutional online feedback platform
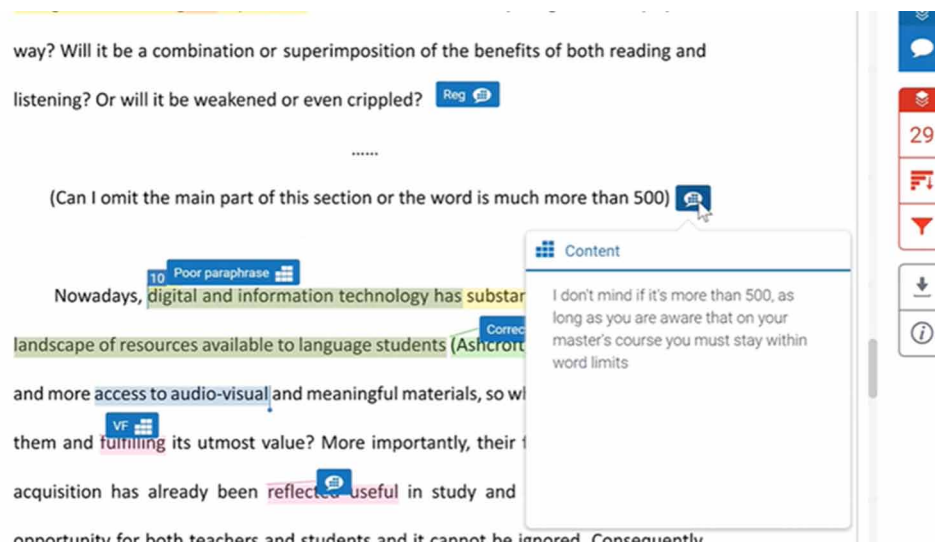
Even though most of the above appears unrelated to S4's use of online rubrics, it provides an insight into some of the reasons why this non-use occurred. Similar to most participants, S4 did not make use of the online rubric in her first think-aloud, even though the feedback was rubric-referenced. Yet, in follow-up interviews, she reported awareness of the rubric: "I have noticed it and I click it on". S4 also commented with confidence on rubrics' role in helping her make sense of grades, "so I can check carefully about the kind of structure, vocabulary, references, what's the grade", and on their potential developmental impact in facilitating her understanding of academic expectations pre-submission and reflecting on own performance post-submission:

I need to look at it and research it carefully before the writing, not just before the feedback… because I need to be familiar with the criteria... And I need to know what level it is about my writing assessment and I can check which part I can improve.

In her final interview, S4 also elaborated on the way rubric score bands have helped her extend her understanding of UK grades, which are notoriously "polysemic ...signify[ing] different things to different students" (Sutton, 2012, p.34): "My final mark is 62. Actually, because it's my first- time encounter with the formal essay… I'm quite satisfied with this final grade… I checked the rubric, it belonged to the 'Merit', so I think it's OK". The rubric appeared to have had positive impact on S4's expectations and modulated the student's affective response to the grade – which is otherwise just a decontextualised number.

And yet, despite providing this additional context for S4, the online aspect seems to have prevented her from engaging with one of the learning opportunities of rubrics: that they break up an overall grade by showing achievement levels for each criterion. In her final interview, for instance, S4 literally enacted the missed encounters with these levels, opening the rubric but not scrolling enough to see that they had been highlighted. The cluttered layout prevented S4 from seeing the various achievement levels, and it was the interviewer who showed S4 how to find them - a moment in the interview that became a metacognitive experience:

Interviewer: *So, when you click here, where it says, "View rubric".*
Yes, I found it.
Interviewer: *And you should see a table with criteria...*
Yeah.
Interviewer: *OK. And you see on the left-hand side, you should see the criteria.*
Yes.
Interviewer: *And is there something about references? One of the criteria should be "Sources" ... or "Support". And then if you look on the right-hand side, you will see the tutor probably has clicked one of them squares that describes the level; it could be "Merit" or "Distinction" but for each criterion you'll be given a level...*
Yes, let me try. Here it is.
Interviewer: *So, see on the left-hand side you have "Sources", yeah? So, if you now use the glide and go further to the right. You know, this glide at the bottom of the page? Keep going. There we go. Oops, you missed it... A bit more to the left. There we go; so, you see the blue [square]? So, for "Sources" you were in the area of 62.*
OK, I got it. I didn't know that before … So, total score is a combination of these sections, right?

This somewhat laborious process of finding the highlighted achievement levels shows what might have prevented S4 from developing further her metacognitive contextual knowledge of online rubrics and grading systems in the UK, and even her monitoring skills. Not unlike participants S1 and S3, S4 shared that the static rubric had been useful and accessible. And, even though S4 repeatedly declared awareness of the online rubric and was willing to learn from the interviewer how to make the most of it, she regretted its "invisible" online aspects: "I didn't notice that before; that's a pity."

Paradoxically, then, in S4's case, it was the onlineness of the rubrics that failed to advance her metacognitive knowledge of her academic context. As a pro-active student who appreciated feedback and

was keen to develop her assessment and feedback literacy, S4 seemed to have found the combination of a static rubric and an alternative learning environment sufficiently accessible, interactive and dialogic. Aware of the online rubric, S4's use of it was limited to seeing it as a static list of criteria to adhere to and a useful code for unpacking the meaning of a numeric grade. She never made use, and thus, sense of the individual score levels within the online rubric, which make up a grade, or noticed its hyperlinking functionalities. By her own admission, the online rubrics could have potentially had more impact on her metacognitive understanding of UK grading practices and her self-regulation, had they been easier to access and navigate. However, when it comes to the kind of dialogic and affective feedback that S4 appreciated, and the student-centred learning environment that she found most productive, it is doubtful that the current online feedback platform could have met her cognitive and metacognitive needs.

S5 CASE STUDY: CONFIDENT RUBRIC USE WITH METACOGNITIVE GROWTH

S5 was a thoughtful and observant student, who continued to do a PhD after the completion of his MSc. He was committed to the project, enthusiastically taking part in a range of dissemination events. S5 was the participant who made the most of the interactive functionalities of the online rubric, using hyperlinks to make sense of feedback and engaging holistically with all elements of the online feedback space. In his metacognitive journey, S5 was supported by the online rubrics into developing a better sense of feedback practices in the UK and a more balanced sense of himself as a learner. His extensive, integrated and attentive use of the online rubric, though, also led him to conclude that a decontextualised online rubric is insufficient for meaningful self-monitoring and effective learning strategies.

S5 approached the understanding of his learner self with a methodical mindset from the very start, although he tended to draw on a deficit framework, a prism through which EAP students are often viewed (Benesch, 2001), saying for example, "my biggest problem about my content is that I showed very little critical thinking". At the same time, he gave considered analysis of the strategies that he might deploy to address the identified deficit, and was able to articulate these in relation to his own context rather than generically, which tended to be the approach of other students:

I think in my future work, when I make some notes, I should give more my own idea... I had read a lot in our classes; it's all about my major. But I just... copied information from other sources... If I'm not critical thinking about it, they're just information, not my knowledge. So... if I want to get some skills in my future work, I have to do some critical thinking process... Such as when I'm reading one paper, in the past, I just think the researcher said very good and always right; but now I will think if I did this work, what method I would choose, and the results would have been better or not compared with this author.

This pronounced procedural orientation was accompanied by a significant degree of confidence when S5 encountered tutor feedback. In contrast to other participants, in his first screencast think-aloud, S5 summarised and sequenced his tutor's overall comments in an order that made sense to him as a learner, instead of reading them verbatim or in a linear fashion. This was particularly evident in his integrated approach to the online feedback space, where he would explicitly draw links between his tutor's overall and in-text comments: "after I read the right window information, I would immediately think of my writing structure in my essay... then, I will go to the left window and... find the details they give me". In his first screencast, S5 highlighted sections of the overall comment in the right-hand window, whilst

summarising the key messages in his own words and then moving to relevant illustrations of the point in his script on the left (see *Figure 7*). S5's approach therefore contrasted with the practices of S1, S2, and S3, who tended to engage with the different elements of online feedback in isolation and sequentially.

Figure 7. S5's integrated use of the online feedback space
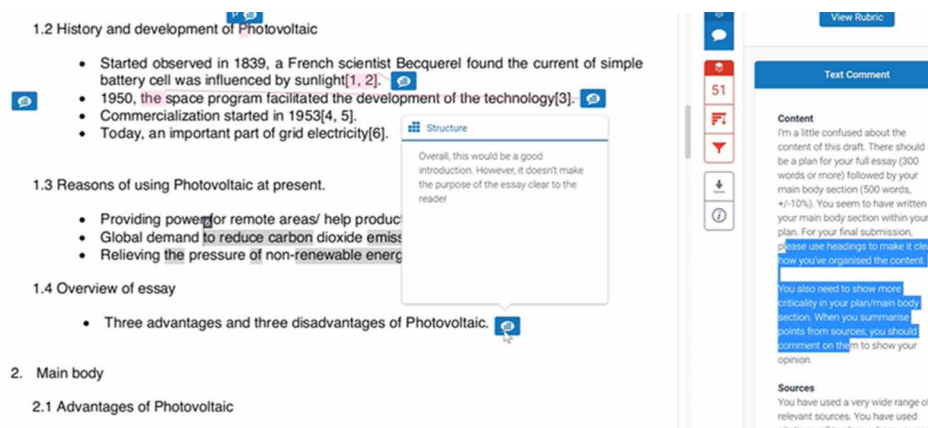
This was sustained when S5 reviewed in-text comments on his script. He not only identified a specific example of a challenge highlighted in the tutor's overall comment, e.g., "lack of clarity", but then proceeded to expand on a related in-text comment by drawing on his subject domain knowledge and his pre-sessional context:

And there are also some details… that giving more unclear information. For example, in the introduction part, I wrote "In 1950 the Space program facilitated the development of the technology." My tutor asked me "Which space program?" because there are various program in different countries. And this question also remind me of the past class with [tutor's name]; she said, "When you write your article, you should make sure your work clear and logical." In other words, the reader of my work shouldn't need to infer my ideas.

Given that in the institutional feedback platform it is not possible for the tutor's overall comments to be hyperlinked to individual in-text comments, unlike the latter being hyperlinked to the online rubric criteria, the think-aloud also demonstrates S5's assured cognitive and metacognitive understanding of feedback comments and academic literacies.

S5's integrated approach to feedback was also evident in his online rubric use. Even though the participant did not make use of the rubric in his first screencast think-aloud, in the follow-up interview it became clear that he was aware of the tool and understood it as codifying performance standards: "My tutor give us one form about the rubric in the beginning of our course… She said we should follow these rubric criteria in future". And yet, unlike participant S4 who had made similar assertions but did not use the online rubric to its full potential, S5 showed fairly in-depth understanding of how the interactive functionalities of the online rubric facilitated his in-task synthesis of multiple feedback elements:

[Y]ou can see in my assignment draft, there a lot of [in-text comments]; so, you know, it looked very complex at first. But after I go to the rubric, I found that there are only four or five information because some information can be classified into the same topic, such as "Structure" ...And that rubric table me get a better understanding my tutor's advice… If I just... read [in-text comments] in my draft, I really didn't know which one… I should have read first and read next because, you know, it's so complex and so difficult.

S5's encounter with the sheer number and complexity of in-text comments shows the additional cognitive load associated with excessive quantity of feedback, which might even result in confusion or poorer performance (Hattie and Timperley, 2007). And in S5's case, the tabular format and hyperlinking functionality of online rubrics facilitated the processing of feedback in a way that made sense to this procedurally oriented student. What can also be seen in S5's second screencast is that his reliance on online rubrics to help him decode the in-text comments determined his navigation of the online feedback space: whereas other participants would first look at the tutor overall comment or in-text comments and end with the online rubric, S5 would use the online rubric earlier to organise his understanding of the in-text comments and refer to the tutor overall comment as a conclusion to his feedback reflection in the think-aloud.

Even though the online rubric supported S5's categorisation of in-text comments and their mapping onto assessment criteria, the inevitable "fuzziness" (Sadler, 1989) of the latter presented a barrier for his further use of this tool for self-regulation. S5 was unsure about the identical wording of rubric descriptors within the same score level: "I have some problem about the scores for 68 and 65 because their details are same with each other. So which type of writing belong to 68?", as well as the nebulous language of the descriptors of the top Distinction level: "Besides how to understand... the word 'highly' at 70?". The metacognitive experience of confusion about the meaning of rubric descriptors speaks to recurrent challenges in the articulation (by teachers) and internalisation (by students) of tacit knowledge within assessment criteria in general (O'Donovan *et al.*, 2004; Orsmond *et al.,* 2002; Popham, 1997). Whilst not necessarily specific to online rubrics, the lack of clarity in the descriptors prompted S5 to consider the role of context in facilitating his cognitive processing of criteria:

But I think there's a problem, you know. If you can give us more examples… about this rubric… I didn't know what's the really meaning of it. I can get a better understanding of it so, if you give me one or two examples about this criteria.

S5's request for "more examples" (i.e., exemplars of student work at different performance levels) suggests that it is not more transparency (see Bearman and Ajjawi, 2021), or granularity (see O'Donovan *et al.,* 2004) in the articulation of rubric criteria that was necessary. Rather, what could have advanced his understanding of rubrics and support his self-regulatory processes was more context and extensive engagement with a range of exemplars as also confirmed in rubric literature (Orsmond *et al.*, 2002; Sadler, 2009).

An interesting metacognitive development in S5's use of online rubrics was the role they had in the adjustment of his understanding of his learner self. In his second think-aloud, S5 noticed that there was no straightforward correlation between the number of in-text comments hyperlinked to a specific criterion and the corresponding achievement level. He noted that the number of in-text comments that appeared in a drop-down box over the "Structure and argument" criterion (5), for instance, exceeded the number of

in-text comments associated with the "Cohesion" criterion (2) (see *Figure 8*), but this did not translate into a lower grade for "Structure and argument": "My tutor give me more details [*in-text comments*] at 'Structure and argument'; but most of these are good thing in my task… But my 'Cohesion' part had more problem". S5's observation that most of the in-text comments "are good" challenged the deficit framework through which he tended to interpret his learner self initially; and in the final interview, this self-knowledge shifted to a more balanced profile – a shift potentially signalled by the rubric itself: "I found most of my skills were good in this rubric".

Figure 8. S5's online rubric with five hyperlinked in-text comments for "Structure and Argument" visible in a drop-down box

Despite its potentially positive impact on S5's self-knowledge, the online rubric had a less clear role in enhancing S5's procedural knowledge. For him, what was missing from the online rubric was the context of the full student script and the guidance of the tutor – perhaps not dissimilar to his earlier request for more contextually relevant exemplars:

I have found grading about my skills in this place [the online rubric]. And the first that I did whether my skill is good or not good. And …I found most of my skills were good in this rubric. But the next step is how to correct these problems and make more improvement. So, I find more details in the next here marks, you know, but I couldn't find the answer over here [the online rubric] …I think the network [hyperlinking] is useful for me because I know clearly which part had more feedback…, but the details on one sentence I think I can't understand them because as I'm reading the sentence I think I need to back to the contents [the script]… I think the overview comments in the [right] window is very important to me because after reading every sentence I find the details in the left window, not in the rubric part.

For S5 then, the online rubric had a role to play in building a balanced view of himself as a learner, a deeper understanding of the UK academic context, and a starting point for his self-regulation. However, in terms of developing appropriate and effective learning strategies, S5 discovered that he "couldn't find the answer" in a rubric space, albeit with hyperlinked in-text comments. His use of the online ru-

bric showed that it was a "productive space" for summative purposes post-submission as it enabled the integration of different online feedback spaces, but this was lacking for developmental purposes due to its insufficient context.

CONCLUSION

Our findings build a complex picture of the range of metacognitive experiences triggered by online rubric use. Being able to see how online rubrics, and the broader online feedback journey of which they are part, are experienced from five different students' perspectives is we feel an inherently valuable step in helping educators understand the challenges and the possibilities for students in this environment. While it is not the purpose of a case study approach to deduce generalizable truths, these five cases do share overlapping elements that can move us to a better understanding of online rubrics' impact on metacognition. The most obvious to note is that of the five cases, only one (S5) was able to engage with the online rubric well, and to exploit the way the in-text comments were hyperlinked to rubric criteria to further develop his already quite high-level metacognition. The other students were hampered in one way or another by the barriers of the online rubrics, for example: they did not initially see how to access them (S2, S3); once accessed, they did not always see the full rubric (S3, S4); and there was no way to engage holistically with the rubric feedback along with the other feedback and feedback from other tasks, requiring students to create their own compiled feedback spaces (S1, S4). From a digital perspective, the separate and work-obscuring way the rubric had to be viewed, combined with it being non-printable, non-shareable, accessible only post-submission, and in a one-way rather than dialogic communication space meant that ultimately the "online-ness" prevented metacognitive development.

However, the students were able to largely overcome these barriers, albeit noting experiences of confusion and frustration with the online rubric, and took differing metacognitive journeys over time to make rubrics work for them. As most of our participants started with an already well-developed metacognitive base, most of them sought to build on the self and contextual knowledge they already had by using print or static electronic rubrics to, for example, develop awareness of tutor expectations and task requirements, use criteria for strategy prioritization, and reflect and self-evaluate. Most participants' encounters with online rubrics prompted reflections on online rubrics' potential to support further the monitoring of learning (e.g., through cross-task feedback records or interactive portfolio) and evaluation of own progress (e.g., through interactive self-assessment activities, peer feedback sharing). Some encounters with the online affordances of rubrics resulted in recognition of their ability to synthesise different levels of external feedback and directed learners to a more in-depth exploration of their contextual knowledge, particularly criteria, grading systems, and feedback. In particular, for S1 and S5, the limitations of online rubrics (inability to see all at once, decontextualised hyperlinked comments) paradoxically prompted further metacognitive development than would have otherwise been. And, with the exception of S2, who developed a minimal rubric use over time that was simply grade and performance-level focused, all the participants developed an understanding and appreciation how rubrics overall (not just online) could help them develop as autonomous learners: as S3 stated in her final interview, "the tutor is the guide; the rubric is the way".

ACKNOWLEDGEMENT

REFERENCES

Andrade, H., & Du, Y. (2005). Student perspectives on rubric-referenced assessment. *Practical Assessment, Research & Evaluation*, *10*(3), 1–11.

Anglin, L., Anglin, K., Schumann, P. L., & Kaliski, J. A. (2008). Improving the efficiency and effectiveness of grading through the use of computer-assisted grading rubrics. *Decision Sciences Journal of Innovative Education*, *6*(1), 51–73. doi:10.1111/j.1540-4609.2007.00153.x

Ashton, S., & Davies, R. S. (2015). Using scaffolded rubrics to improve peer assessment in a MOOC writing course. *Distance Education*, *36*(3), 312–334. doi:10.1080/01587919.2015.1081733

Atkinson, D., & Lim, S. L. (2013). Improving assessment processes in Higher Education: Student and teacher perceptions of the effectiveness of a rubric embedded in a LMS. *Australasian Journal of Educational Technology*, *29*(5), 651–666. doi:10.14742/ajet.526

Bearman, M., & Ajjawi, R. (2021). Can a rubric do more than be transparent? Invitation as a new metaphor for assessment criteria. *Studies in Higher Education*, *46*(2), 359–368. doi:10.1080/03075079.2019.1637842

Benesch, S. (2001). *Critical English for Academic Purposes: theory, politics and practice*. Routledge. doi:10.4324/9781410601803

Bloxham, S., & Campbell, L. (2010). Generating dialogue in assessment feedback: Exploring the use of interactive cover sheets. *Assessment & Evaluation in Higher Education*, *35*(3), 291–300. doi:10.1080/02602931003650045

Bowles, M. (2010). *The think-aloud controversy*. Routledge.

Brookhart, S. M., & Chen, F. (2015). The quality and effectiveness of descriptive rubrics. *Educational Review*, *67*(3), 343–368. doi:10.1080/00131911.2014.929565

Butler, D. L., & Winne, P. H. (1995). Feedback and self-regulated learning: A theoretical synthesis. *Review of Educational Research*, *65*(3), 245–281. doi:10.3102/00346543065003245

Carless, D., & Boud, D. (2018). The development of student feedback literacy: Enabling uptake of feedback. *Assessment & Evaluation in Higher Education*, *43*(8), 1315–1325. doi:10.1080/02602938.2018.1463354

Creswell, J. W., & Creswell, J. D. (2023). *Research design: qualitative, quantitative, and mixed methods approaches* (6th ed.). SAGE.

Deci, E. L., & Ryan, R. M. (1985). *Intrinsic motivation and self-determination in human behavior*. Plenum Press. doi:10.1007/978-1-4899-2271-7

Dornyei, Z. (2007). *Research methods in applied linguistics: quantitative qualitative, and mixed methodologies*. Oxford University Press.

Ericsson, K., & Simon, H. (1993). *Protocol analysis: verbal reports as data* (2nd ed.). MIT Press. doi:10.7551/mitpress/5657.001.0001

Flavell, J. H. (1979). Metacognition and cognitive monitoring: A new area of cognitive–developmental inquiry. *The American Psychologist*, *34*(10), 906–911. doi:10.1037/0003-066X.34.10.906

Gall, M. D., Gall, J. P., & Borg, W. T. (2003). *Educational research* (7th ed.). Pearson Education.

Gass, S. M., & Mackey, A. (2016). *Simulated recall methodology in applied linguistics and L2 research* (2nd ed.). Routledge. doi:10.4324/9781315813349

Hattie, J., & Timperley, H. (2007). The power of feedback. *Review of Educational Research*, *77*(1), 81–112. doi:10.3102/003465430298487

Haworth, R. (2016). Personal Learning Environments: A solution for self-directed learners. *TechTrends*, *60*(4), 359–364. doi:10.100711528-016-0074-z

Hyland, F. (1998). The impact of teacher written feedback on individual writers. *Journal of Second Language Writing*, *7*(3), 255–286. doi:10.1016/S1060-3743(98)90017-0

Jonsson, A., & Svingby, G. (2007). The use of scoring rubrics: Reliability, validity and educational consequences. *Educational Research Review*, *2*(2), 130–144. doi:10.1016/j.edurev.2007.05.002

Krebs, R., Rothstein, B., & Roelle, J. (2022). Rubrics enhance accuracy and reduce cognitive load in self-assessment. *Metacognition and Learning*, *17*(2), 627–650. doi:10.100711409-022-09302-1

Leki, I. (1995). Coping strategies of ESL students in writing tasks across the curriculum. *TESOL Quarterly*, *29*(2), 235–260. doi:10.2307/3587624

Leki, I. (2001). "A narrow thinking system": nonnative-English-speaking students in group projects across the curriculum. *TESOL Quarterly*, *35*(1), 39–67. doi:10.2307/3587859

McKinney, B. (2018). The impact of program-wide discussion board grading rubrics on students and faculty satisfaction. *Online Learning : the Official Journal of the Online Learning Consortium*, *22*(2), 289–300. doi:10.24059/olj.v22i2.1386

Nicol, D. J., & Macfarlane-Dick, D. (2006). Formative assessment and self-regulated learning: A model and seven principles of good feedback practice. *Studies in Higher Education*, *31*(2), 199–218. doi:10.1080/03075070600572090

Nordrum, L., Evans, K., & Gustafsson, M. (2013). Comparing student learning experiences of in-text commentary and rubric-articulated feedback: Strategies for formative assessment. *Assessment & Evaluation in Higher Education*, *38*(8), 919–940. doi:10.1080/02602938.2012.758229

O'Donovan, B., Price, M., & Rust, C. (2004). Know what I mean? Enhancing student understanding of assessment standards and criteria. *Teaching in Higher Education*, *9*(3), 325–335. doi:10.1080/1356251042000216642

Orsmond, P., Merry, S., & Reiling, K. (2002). The use of formative feedback when using student derived marking criteria in peer and self-assessment. *Assessment & Evaluation in Higher Education*, *27*(4), 309–323. doi:10.1080/0260293022000001337

Panadero, E., & Jonsson, A. (2013). The use of scoring rubrics for formative assessment purposes revisited: A review. *Educational Research Review*, *9*, 129–144. doi:10.1016/j.edurev.2013.01.002

Panadero, E., & Jonsson, A. (2020). A critical review of the arguments against the use of rubrics. *Educational Research Review*, *30*, 1–19. doi:10.1016/j.edurev.2020.100329

Pekrun, R., Goetz, T., Titz, W., & Perry, R. P. (2002). Academic emotions in students' self-regulated learning and achievement: A program of qualitative and quantitative research. *Educational Psychologist*, *37*(2), 91–105. doi:10.1207/S15326985EP3702_4

Pintrich, P. R. (2002). The role of metacognitive knowledge in learning, teaching, and assessing. *Theory into Practice*, *41*(4), 219–225. doi:10.120715430421tip4104_3

Pitt, E., & Norton, L. (2017). "Now that's the feedback I want!": Students' reactions to feedback on graded work and what they do with it. *Assessment & Evaluation in Higher Education*, *42*(4), 499–516. doi:10.1080/02602938.2016.1142500

Popham, W. J. (1997). What's wrong – and what's right – with rubrics. *Educational Leadership*, *55*, 72–75.

Reed, P., Watmough, S., & Duvall, P. (2015). Assessment analytics using Turnitin and Grademark in an undergraduate medical curriculum. *Journal of Perspectives in Applied Academic Practice*, *3*(2), 92–108. doi:10.14297/jpaap.v3i2.159

Rhodes, M. G. (2019). Metacognition. *Teaching of Psychology*, *46*(2), 168–175. doi:10.1177/0098628319834381

Robbins, J., & Marinkova, M. (in press). Students' (non)use of online rubrics: Turnitin vs feedback literacy. *Practitioner Research in Higher Education*.

Ryan, J., & Viete, R. (2009). Respectful interactions: Learning with international students in the English-speaking academy. *Teaching in Higher Education*, *14*(3), 303–314. doi:10.1080/13562510902898866

Ryan, R. M., & Deci, E. L. (2000). Intrinsic and extrinsic motivations: Classic definitions and new directions. *Contemporary Educational Psychology*, *25*(1), 54–67. doi:10.1006/ceps.1999.1020 PMID:10620381

Sadler, D. R. (1989). Formative assessment and the design of instructional systems. *Instructional Science*, *18*(2), 119–144. doi:10.1007/BF00117714

Sadler, D. R. (2009). Indeterminacy in the use of preset criteria for assessment and grading. *Assessment & Evaluation in Higher Education*, *34*(2), 159–179. doi:10.1080/02602930801956059

Sadler, D. R. (2014). The futility of attempting to codify academic achievement standards. *Higher Education*, *67*(3), 273–288. doi:10.100710734-013-9649-1

Schraw, G. (1998). Promoting general metacognitive awareness. *Instructional Science*, *26*(1/2), 113–125. doi:10.1023/A:1003044231033

Suskie, L. (2017). Rubric Development. In C. Secolsky & D. B. Denison (Eds.), *Handbook on measurement, assessment, and evaluation in Higher Education* (2nd ed., pp. 545–559). Routledge. doi:10.4324/9781315709307-43

Sutton, P. (2012). Conceptualizing feedback literacy: Knowing, being, and acting. *Innovations in Education and Teaching International*, *49*(1), 31–40. doi:10.1080/14703297.2012.647781

Torrance, H. (2007). Assessment as learning? How the use of explicit learning objectives, assessment criteria and feedback in post-secondary education and training can come to dominate learning. *Assessment in Education: Principles, Policy & Practice*, *14*(3), 281–294. doi:10.1080/09695940701591867

Turnitin. (2021). *A New Path and Purpose for Turnitin*. Turnitin. https://www.turnitin.com/blog/a-new-path-and-purpose-for-turnitin

Wenden, A. (1987). Metacognition: An expanded view of the cognitive abilities of L2 learners. *Language Learning*, *37*(4), 573–594. doi:10.1111/j.1467-1770.1987.tb00585.x

Winstone, N., & Carless, D. (2022). *Designing effective feedback processes in Higher Education: a learning-focused approach*. Routledge.

Yang, M., & Carless, D. (2013). The feedback triangle and the enhancement of dialogic feedback processes. *Teaching in Higher Education*, *18*(3), 285–297. doi:10.1080/13562517.2012.719154

Young, P. (2000). "I might as well give up": Self-esteem and mature students' feelings about feedback on assignments. *Journal of Further and Higher Education*, *24*(3), 409–418. doi:10.1080/030987700750022325

Zimmerman, B. J. (1990). Self-regulated learning and academic achievement: An overview. *Educational Psychologist*, *25*(1), 3–17. doi:10.120715326985ep2501_2

Chapter 15
Exploring and Developing Reflective Writing Rubrics in Higher Education

Martin Sands

King's College London, UK

ABSTRACT

This chapter highlights key academic contentions around assessing reflection in higher education. Through assimilating experiential, conceptual, and empirical evidence, key terms are defined: 'reflection' as a conceptual and neuroscientific phenomenon within higher education; 'reflective practices' as activities, models, and frameworks that facilitate it; and 'reflective assessments' as content or process driven products of those practices. A literature review elucidates three voices that speak to reflection in higher education.

INTRODUCTION: CHARACTERISING REFLECTION IN HIGHER EDUCATION

Before designing learning activities and assessments that seek to assess reflection, it is important to contextualise its existence and facets that are the potential focus of assessment. Reflection is a phenomenon in higher education conceptualized upon theories of adult education, yet its component parts, function and constructs are varied and abstract. In his concept analysis for higher education Rogers (2001) deconstructs the phenomenon of reflection through synthesizing some fifteen different and heterogenous theories. Of note, Rogers (2001) describes the process, which encapsulates the steps involved in the reflective thought or reflective practice. This includes: 1.) identifying a problem or event that needs addressing; 2.) collecting further information and an awareness of different perspectives, questions and assumptions; 3.) deciding to act including reframing, meaning making and transforming previous understandings; 4.) taking action related to the previous steps. Rogers (2001) also describes the outcomes which includes the learning from the reflective practice, which are demonstrated through different assessment modes (including written and oral means). Reflective assessments, therefore, involve several factors for consideration, including whether the assessment is content driven (focused on outputs and

learning), or process driven (focused on how an individual got to their new learning), or both. Despite commonalities identified in the various concepts of reflection (Rogers, 2001), numerous authors have debated the heterogeneity of its inputs and outputs and how these interact with educational activities and assessment. It follows then, that when designing learning reflective assessments and activities, it is important to understand the key voices, contentions and conventions related to reflection. It is also important to consider how each impact on meeting learning outcomes, inclusive learning, and assessment fairness. A systematic approach towards searching the literature has identified opposing opinions about reflection, reflective assessments and reflective practices culminating in three different philosophical viewpoints or *voices* that speak to reflection: the *reductionist*, *pluralist* and *pragmatic*. Ng et al. (2015) introduced the concepts of reductionist and pluralist elements of reflection, whereas the third voice, the pragmatic, is introduced here as it emerges from the literature through tensions between the other two.

While a systematic approach to the literature has cast a wide net, much of what emerged was found in medical and healthcare related subjects including medicine, physiotherapy, nursing, and pharmacy, with a smaller proportion in social sciences and teaching. Given the common educational theories, it is hoped that the literature in this chapter provides a relevant or parallel context for other professions too.

Reflection: Reductionist Entity or Pluralist Phenomenon

In designing criteria and descriptors for different modes of assessment, it is important to understand what the assessment mode looks like. For example, practical exams, laboratory reports, presentations and exams usually have some defining structure by which a learner and assessor can mutually agree a scaffold that allows a learner to demonstrate outcomes of learning. However, despite the integration of reflection in higher education, its definition and defining characteristics remain somewhat an enigma, with more work needed to provide an agreement on its definition and processes by students and educators alike (Rogers, 2001; Uygur et al., 2019). Historically, numerous seminal authors have defined reflection theoretically or conceptually (Dewey, 2006; Kolb, 1984; Schön, 1983), and a review of past and contemporary literature has found two strongly opposing voices within the reflective debate: the *reductionist* and *pluralist* .

The *reductionist* voice speaks of seeking a definition of reflection, and a way to prove its existence as an independent entity, presumably to help ring-fence the teaching of reflection and its process to developing professionals. In the 1990's, concurrent to the rapid inclusion of reflective assessments in professional programmes, (Johns & Freshwater, 1998, p. 2) remarked, 'It seems an academic pastime to define exactly what it *[reflection]* is.' Since then, integration of reflection and reflective assessments in curricula has grown across a broad spectrum of higher education programmes. Reflective frameworks and models are regularly cited as nurturing personal attributes such as empathy, emotional intelligence, self-awareness and resilience, and are included in guidance documents for the completion of reflective essays and professional portfolios. Written reflective statements are used to support the validation and revalidation of licenses for practice professionals, and the skill of reflection is regularly highlighted as contingent to the effective practitioner. Over time, reflection and reflective practices have adopted increasingly higher stakes functions in professional education. Literature since John's and Freshwaters' observations has demonstrated ongoing attempts to legitimise the role of reflection in professional and academic practice. Perhaps reducing reflection to its constituent parts has become less of a 'pastime', and more of an essential preoccupation to align its educational purposes to its intended outcomes.

However, while the *reductionist* voice seeks to provide a definition to guide educational activities, other authors have suggested that defining reflection is incompatible with its phenomenological origins and intent. Supporting the *pluralist* voice, Ng et al. (2015, p. 461) debate whether reflection can be simplified to a 'utilitarian' competency-based tool when this sits inharmoniously with its origins as a complex positionality, or way of being in and seeing the world. This debate is not new. In nursing, Rolfe & Gardner (2006, p. 24) suggested defining reflection oversimplifies its origins as a unique and personalised learning phenomenon towards a dichotomous 'technical tool' with the unhelpful connotation that there is a right and a wrong way.

Anecdotal evidence in higher education sees the *reductionist* voice speak loudest. The mission to define the premises and processes of reflection continues, aiming to design curricula that can teach and assess it. Reductionism has, ironically, spawned a heterogeneity of research that informs the processes in teaching and assessments (Uygur et al., 2019), dependent on who has defined it, and how it has been separated into its constituent characteristics. Whilst the *pluralist* voice states that reflection is not a simple definable entity, anecdotal evidence demonstrates that a *reductionist* paradigm for assessing reflection remains pervasive in higher education, with a many reflective assessment modes including journals, essays and viva voce where learners are encouraged to think-aloud. Assessment criteria and rubrics are often applied to these activities, but it is interesting to consider the drivers and science that informs how reflection is assessed.

Neuroscience and Reflective Zombies

Attempts to identify the science of reflection, seated in social cognitive neuroscience, harmonize the reductionist voices researching reflection. While certain brain structures have been associated with activities deemed reflective, there appears to be a more complex neuro-matrix involved in the production of reflective thought (Lieberman, 2012). Lieberman (2012, p. 294) explores reflective thought by comparing the reflective human mind to the metaphorical, and non-reflective, zombie mind. Lieberman describes structures involved in two hypothetical systems, with the C and X systems involved in reflective and non-reflective thinking respectively. Lieberman's overview of structures involved in different types of thought are briefly described below (see table 1).

Table 1. Lieberman's (2012) conceptualization of the systems of reflective and non-reflective consciousness

System	Characteristics	Points of difference
X-system (includes amygdala, ventromedial prefrontal cortex, dorsal anterior cingulate cortex, lateral temporal cortex)	Spontaneous Functions in parallel to other brain functions Fast acting reactions Triggered by high arousal Reacts to common events in external world	Non-reflective Maintains attention on the external world and real events as they happen.
C- system (includes lateral prefrontal cortex, rostral anterior cingulate cortex, medial temporal lobe)	Intentional, calculated Self-generated outputs run in series to each other Slow acting responses Inhibited by high arousal Responds to exceptional cases and events through thinking about the internal world	Reflective Needs space, time, quiet and an exceptional case to which to apply reasoning and logic. Directs attention towards internal world.

To compare the two systems, the non-reflective X-system of consciousness characterises the 'zombie' mind with thinking that is spontaneous to the external world and reacts to high levels of sensory arousal. In contrast, the reflective C-system of consciousness goes beyond simple reactions. Lieberman describes reflective thought as further thinking upon experiences which form part of our previous flow of consciousness; a thinking that asks further questions about what happened. Thinking upon thinking, or metacognition, allows us to respond to sensory arousal by using situational judgement, self-inquisition and a focus on linguistics. This reflective consciousness or C-system allows us to deliberately delve into our internal world and calculate selective responses to past events. According to Lieberman, the reflective C-system differs from the non-reflective X-system in that it processes an additional meta-cognitive load, which leads to selective changes in behaviour. This concurs with theoretical definitions of reflection that describe a purposeful focus on change, improvement, and in more practical terms, maintaining good practice or preventing the same error twice.

Understanding the neuroscience prompts critical questions about the aims of reflective assessments. With the reductionist project focusing on assessing a process of reflection, concerns arise that students are led to use the X-system, reacting to assessment checklists without really responding reflectively at all. On this premise, De La Croix & Veen (2018, p. 394) introduce the 'reflective zombie', a learner who displays all the outer signs of having done reflection (e.g., the steps for a reflective essay or assessment), without truly reflecting on internal experiences at all. De La Croix & Veen (2018) amplify the pluralist voice stating that measuring reflective process is detrimental to its learning impact. They oppose reductionist assessments that drive learners towards a task-based process for 'doing' reflection rather than engaging the learner in the content of learning it is supposed to produce. In assessment terms, someone might be seen to reflect when judged against a process, but this might not represent any meaningful learning output for the learner or end-user of their practice.

It is interesting that the zombie metaphor appears in both theoretical educational and empirical cognitive neuroscience literature around reflection, and the differentiated X and C systems provide a sense that a 'reflective zombie' could exist. However, Lieberman's exploration does not fully support this. While the C and X system both appear to be involved in reflective thought, it is not yet fully known whether one system is preceded, inhibited, or attenuated by another during reflective thought, or whether the two types of thought are indeed separable. Activities requiring meaningful and responsive learning can be mapped, at least crudely, to a set of regions in the brain. However, the complexity of cognitive neuroscience and the seemingly intangible phenomenon of reflection itself makes it reasonable to assume that isolating regions of the brain designated to reflective learning would need a lot more complicated research.

So, we are some way off using functional MRI scanners to discover whether learners have really reflected or not, and we are unlikely to see such assessment measures being implemented in higher education any time soon! Yet, commonalities in both educational and neuroscience literature do point to certain essential requirements for reflective thought and reflection to occur. Reflection and reflective thought require calculated and deliberate responses to events, rather than simply completing rudimentary and reactive steps towards a process of reflection. Reflection supports, and is supported by, behaviours invested in meaningful learning including time, metacognitive effort and a non-distracting space to think about specific triggering events. As such, reflective assessment briefs should encourage learners to make time and effort to think about their own thinking and ask them to create action plans and reconceptualisations that demonstrate considered responses to specific events.

Empirico-Professional Context

The limited evidence for reflection as a science means the existence of reflective assessments must be strongly reinforced by other influences. Regulatory and statutory bodies have adopted reflection as a vehicle for learning and stipulate that regular reflection is an essential proficiency and attribute of a qualified practitioner (General Medical Council, 2023; Health and Care Professions Council, 2023). Increasingly, validation and approval of professional education programmes requires proof of measurable, legitimate, and evidence-based interventions that guide learners towards target outcomes. The pervasiveness of the re*ductionist* voice in designing reflective assessments is driven by the need to teach and objectively measure reflection as an essential outcome and professional skill. While regulatory bodies continue to promote the importance of reflection in the development of professional identity, the drive for a cost-effective and legitimising evidence-base continues to justify educational interventions that support professional practice. The portmanteau of an 'empirico-professional' is used here to frame the often reductionist efforts made to define reflection to meet the requirements for professional registration. For example, Donaghy & Morss (2009, p. 6) defined reflection in the context of physiotherapy practice. Using expert consensus, they adapted an existing definition:

"The higher order intellectual and affective activities in which physiotherapists engage to critically analyse and evaluate their experiences in order to lead to new understandings and appreciation of the way they think and operate in the clinical setting."

While this expert consensus provided a framework for learning about person-centred and professional reflection in physiotherapy education in the year 2000, a more robust empirical definition of reflection was lacking, until 14 years later. The field of medical education consistently produces research around reflection and in a systematic literature review, Nguyen et al. (2014) synthesised several definitions of reflection to formulate their own in medical education. They advanced reflection as what a learner thinks about (the content), how they undertake that thinking (the process), and a focus on change (the aims). The final definition (Nguyen et al., 2014, p. 1182) is compact and concise,

'Reflection is the process of engaging the self in attentive, critical, exploratory and iterative inter-actions with one's thoughts and actions, and their underlying conceptual frame, with a view to changing them and with a view on the change itself.'

In medicine, evidence supports reflective writing as a vehicle to develop stronger professional identities, professional values and improve advanced problem solving in the workplace (Lim et al., 2023; Winkel et al., 2017) and perhaps searches for a definition are driven by the importance placed on practices such as reflective writing. Nguyen et al's. (2014) definition remains the most recent conceptual definition founded upon an empirical methodology, and while it provides a set of defining processes on which to base assessment criteria it also gives further resonance to the *reductionist* voice for reflective assessment. The formation of professional identity and values are, of course, unarguable essentials for professional programme delivery, but evidence highlights several areas for caution for assessment design when applying a single definition for 'reflection as an entity' in assessments. These include learner diversity and assessment fairness, and whether providing a definition of reflection alone is in fact helpful.

The Emerging Pragmatic Voice

Research prior to the empirically constructed definition of Nguyen et al. (2014) attenuated the reductionist voice in finding that providing learners with a definition of reflection alone does not help them to reflect, but providing a guiding framework does (Aronson et al., 2012). The positive impact of guiding learners to reflect effectively is a common narrative of a third voice that emerges from the literature around reflection. This voice speaks less of the tension between the reductionist and pluralist voices, and instead focuses on guiding principles that help learners meet the professional and academic requirements of reflective learning and assessment. With empirico-professional contexts of reflection espousing the benefit of reflective practice on professional attributes, a *pragmatic voice,* introduced and advocated here, emerges from the synthesis of tensions in the literature and anecdotal experiences of teaching and assessing students in higher education. Reflective assessments continue to form a large component of measuring professional capability in educational programmes and the pragmatist will want to find a middle ground between fair assessment and learning that authentically and effectively meets professional learning needs. The two opposing voices in reflection are hard to reconcile. As a pluralist practice phenomenon, reflection might be easy to explore, but as a reductionist neuroscientific entity, it might be hard to explain. The pragmatic voice speaks in the grey areas of academic practice, seeking to accept differences and get the job done in the best interests of the learner and the people they will eventually serve.

RUBRICS FOR REFLECTIVE WRITING

Focus and Language of Criteria

Getting the job done has led researchers and authors to adopt a pragmatic approach to reflective assessment in higher education. Aronson et al's (2012) randomised-controlled trial demonstrated that learners do better when provided guidance towards writing a reflective assignment, as opposed to being provided with a definition of reflection alone. This validates past work that has framed effective reflection with guidance around what features of reflection optimise learning, rather than getting distracted by definitions. Rubrics constitute one framework that describe performance and have been used to facilitate learning through reflective writing, with multiple authors describing them as 'road maps' to guide learners' reflections (Devlin et al., 2010, p. 1143; Monbec et al., 2021, p. 10) and scoping reviews have emerged to recommend the use of longitudinal assessment rubrics in reflective writing in medical education (Lim et al., 2023). However, there are some generic assessment limitations that need to be considered before implementing rubrics for reflective assessments.

Long held contentions exist in the literature exploring assessment that drives effective learning. Sadler (2009) discusses the pitfalls of codifying learning into assessment criteria for grading, particularly when the mode, like reflective assessments, is divergent and not subject to a single best or right answer. Torrance (2007) also highlights a generic risk of reductionism in assessment, opposing overly narrow and explicit descriptions of learning performance that drive students to focus only upon meeting criteria to pass assessment. Both authors question assessment of divergent learning upon a narrow set of descriptors where wider learning is intended. Torrance (2007) argues that assessment should not be seen as learning. Instead, it should be the vehicle for learning and/or a measure of learning. This is a significant

concern for educators who wish to drive deep learning from experiences, as in reflective assessment contexts, unintended moves that drive superficial learning to meet assessment criteria give further life to the reflective zombie, with learners missing opportunities for deeper learning from real-life experiences.

Accounting for Torrance's (2007) concerns around assessment *as* learning, helpful rubrics would provide a set broad set of descriptors to help students understand what the reflective assessment requires to reach effective learning. However, formulating guidance includes identifying the parameters that ensure reflective learning occurs. Brookhart (2013) advises against specific task-based language in rubric design, instead encouraging language that describes how learning looks, which sits harmoniously with the neuroscientific basis of reflection. There are also practical limitations in trying to assess the divergent content of various reflections – there are likely to be no 'right' answers. This leaves a space for assessment language that describes the learning performance, as opposed to that which describes the completion of the steps of a reflective process or the content of learning that is applied in practice. Moon's (2007) approach to facilitating learning through reflective writing fits into this space, describing how far reflective writing should dive into experiences through using a 'depth dimension' (Moon, 2007, p. 194). Moon's work includes exemplar versions of 'The Park' which depict the difference between superficial and more critical narrative reflections on events. These exemplars are useful in meeting the pluralist voices asking for educational guidance that shows how depth of learning looks in reflective assessments (De La Croix & Veen, 2018). Such exemplars and frameworks are useful developmental 'yard-sticks'. They scaffold students towards writing in sufficient depth to encourage effective reflective learning from events. Depth is a recurrent metaphor in the discourse surrounding the design of assessment frameworks for reflection and is often aligned to learning (Greenfield et al., 2015; Hatton & Smith, 1995; Moon, 2007; Wald et al., 2012). Research highlights that students often struggle to reach a sufficient depth when asked to write reflectively about experiences (Greenfield et al., 2015). While one could argue that measuring depth does not necessarily translate to applied learning in practice, it does perhaps hold a key to unlocking what effective reflective learning looks like beyond a process-based 'tick box' assessment. The focus on depth moves away from auditing the process or judging the content of reflection, and instead focusses on the quality of learning. Moon maintained that it is not enough to provide a definition of reflection and then expect someone to be able to reflect (Moon, 2007). This is substantiated by Aronson et al. (2012) and provides a pragmatic evidence base that justifies giving formative guidance to learners that promotes depth in reflective writing.

Examples of Existing Evidence-Based Rubrics

The use of rubrics to guide the development of reflective writing is further supported in the research literature, and some include a focus on depth. Sometimes, incorporating this into practice requires a pragmatic acceptance of reductionist research. For example, when Wald et al (2012) developed a rubric to assess reflective writing proficiency, they later adopted a *reductionist* approach to redefine reflection (Wald, 2015), and as such might be seen to portray 'reflection as a learning tool' (Winkel et al., 2017). Such reductionism is unsurprising when empirico-professional influences are considered and educational practice intends to assess reflection as a competence or capability. Expecting teachers to teach and assess something, and learners to learn from something that has no guidance or theoretical framework seems incommensurate with writing clear learning outcomes.

Following these premises, Wald et al. (2012) developed the REFLECT rubric. In line with neuroscientific systems, this analytic rubric focuses on moving students through 4 levels of reflective

writing, from non-reflective towards reflective. It also implicitly alludes to depth, using words such as 'superficial' to describe a lacking depth in non-reflective writing. This rubric has been subject to trials and studies which support its use (Miller-Kuhlmann et al., 2015; Wald et al., 2012). Despite evidence showing variable levels of reliability with the REFLECT rubric (Grierson et al., 2020; Miller-Kuhlmann et al., 2015; Soemantri et al., 2022; Wald et al., 2012) such research is increasingly valuable to inform pragmatic educators on the best ways to get the job done when the absence of a 'gold standard tool' exists (Miller-Kuhlmann et al., 2015). Lucas et al. (2019) also developed and tested a rubric in second year students taking a master's in pharmacy and found that it was a reliable tool for assessing written reflection. However, their rubric is less descriptive of depth and instead focuses on describing how a student engages with feelings and understandings of the self in determining non-reflective, reflective and critically reflective written accounts. Their conclusions give support to the reductionist voice, using a scoring system to rate how learners apply a process of reflection, but also take a pragmatic approach in describing what reflective learning looks like. Monbec et al. (2021) took a combined approach to use process-based and depth-based language in a rubric designed for the development of reflective capability in first year nursing students. Their rubric included sections of Gibbs Cycle (Gibbs, 1988) as a guiding framework for students to bridge the gap between what students and assessors had previously taught and been taught. The final rubric provides guidance for a process of reflection and a description of what the reflective learning looks like.

A point of difference emerges in research that evaluates the NARRA rubric (Alsina et al., 2017). While the aforementioned rubrics have been used to assess and provide feedback to students, what is less clear is whether or how they are used as a student-facing formative aid prior to reflective writing. Alsina et al. (2017), however, not only found the NARRA rubric valid and reliable for measuring the level of reflection in pre-service teachers, but also collected qualitative data demonstrating that most participants found the NARRA useful as a student-facing guide to doing reflection. Guidance in the form a clear student and assessor facing brief and rubric has been found to reduce the ambiguity of assessors' value judgements in reflective assessments (Cheng & Chan, 2019) and sharing rubric development with students is key to finding a common language and expectation. The NARRA rubric projects the reductionist voice through developing a tool for a prescribed process which connects reflection to action. However, the authors also voice a more pragmatic and longitudinal perspective by providing learner-facing guidance for lifelong learning. They further suggest that reflection is not just learnt through teaching but also through lived experience (Alsina et al., 2017). Their philosophy is that a rubric can help prompt making those connections.

Rubrics like these might be advanced as useful learner and educator tools to navigate the requirements of reflective assessments and share some common features. Philosophically, they all sit in a pragmatic space which includes embracing some degree of reductionism in presenting a process or definition for reflection, while going some way to satisfy pluralist perspectives that ask for description rather than prescription of reflective practice (De La Croix & Veen, 2018). Key to these commonalities is the language used and when the opinions of Torrance (2007) and Sadler (2009) are considered, perhaps less important is the type of tool used to facilitate and assess reflective writing than the narrowness of language used to describe performance. Such rubrics have an evidence-base and theoretical foundation by which they can be used or adapted for the development of reflective practices. However, while these rubrics are built upon carefully developed research, they still require further consideration before widespread adoption into a variety of educational contexts. All research comes with limitations, which are further considered here.

LIMITATIONS TO USING REFLECTIVE RUBRICS

Sharing Feelings and Emotions

After the development of the REFLECT rubric, Wald (2015) redefined reflection to try to bring further clarity to the context of reflective writing for medical education. This definition, is partly an expansion on Nguyen et al's (2014) definition:

'Reflection is a metacognitive process including connecting with feelings that occurs before, during, and after situations with the purpose of developing greater awareness and understanding of self, other, and situation, so that future encounters with the situation including ways of being, relating, and doing are informed from previous encounters.'

To note, this definition adds 'feelings' to the mix, citing their importance for professional identity formation in healthcare. On the REFLECT rubric, attending to feelings provides a descriptor of the more competent reflector. This, in contrast with Nguyen et al's. (2014) omission of the words 'affective', 'feelings' and 'emotions' from their definition, allowing a learner to include or exclude this element as they deem appropriate. Higher levels of success in reflective writing are similarly contingent on attending to and sharing feelings in the rubric evaluated by Lucas et al. (2019). Monbec et al (2021) designated significant importance to emotions and feelings in their discussions around the design of a reflective rubric and even included Gibb's Cycle (Gibbs, 1988) which conceptualises feelings as a component of reflection. 'Understanding-Emotions' and 'Personal feelings' also featured as contingent to proficient reflection in a more generic and cross-profession rubric in higher education (Rogers et al., 2019, p. 765). However, the authors here explicitly described a need to separate cognitive and emotional understanding but did not qualify the reasons for this. These sit in contrast with the NARRA rubric (Alsina et al., 2019) which focusses more on an individual's awareness of their own assumptions and beliefs in deeper levels of reflection.

The inclusion of feelings has a positive intention. It leads the learner to think about the more emotional aspects of patient-centred care in healthcare related subjects. Coulehan and Granek (2012) and Monbec et al (2021) develop this further, suggesting that reflection and reflective writing can enable medical professionals to maintain the separation between objectivity and compassion required for effective medical practice. However, there are potential negative consequences to reflective writing and assessment when a focus on sharing emotions and feelings becomes contingent on reflective ability. The pluralist voice questions whether it is a helpful to expect all students to perform equally under the umbrella of reflection, particularly when focusing on the self, thoughts, feelings and emotions at different levels of learning and development (Ng et al., 2014). Students and educators are often perplexed about the extent to which learners are expected to reflect on their own feelings and actions, and how much of this they should share in their reflective assessments (Chan & Lee, 2021). Data from an interrater reliability study for a reflective assessment rubric highlighted that assessors' perceptions of what constituted 'attending to feelings' in written reflections was noticeably different between raters (Lucas et al., 2019). Lucas et al. (2019) found that attending to feelings showed the least inter-rater agreement in their reliability study, perhaps attributed to a lack of consensus on what this criterion describes. Autoethnography has also uncovered levels of anxiety shared by educators and students related to the content and context of

the reflective focus, including the impact of attending to feelings and emotions (Watts, 2015). Learners are not only anxious about the task of a reflective assignment but also affected by developing its content.

The perspective of learner diversity also raises concerns about focusing on emotions in reflective writing. The use of rubrics and models that include a focus on feelings as part of reflection seldom explore how students attend to and deal with subsequent emotions (Monbec et al., 2021). Neurodiversity is an emerging 'hot topic' in high education and learners with preexisting mental health conditions, like autism spectrum disorder, attention deficit and hyperactivity disorder, or those who are transitioning into independent adult life, may be more adversely challenged by the higher education climate and environment (Clouder et al., 2020). Given a reflective assessment's potential to elevate emotional affect, questions arise whether a focus on the emotional aspects of reflective assessments might be too much for some, especially in the formative years of adult education. It is interesting to note that Monbec et al. (2021) observed a lesser focus on transformative aspects of learning in first year nursing students' written reflections. Whether this was due to them using an analytic rubric guide students towards a process of reflection, or whether it was because students were not ready or able to engage in this depth of reflection is unknown. However, when considered alongside the potential risks of heightened emotions, learners might be progressed more safely from instrumental towards emotional aspects of reflective writing through specific rubrics, and as Monbec et al. (2021, p. 10) state, help them explicitly learn the 'rules of the game'.

If a learner is less able or keen to share or engage with difficult emotions, they might fabricate less bothersome events to meet descriptors. Maloney et al. (2013) found fabrication to be particularly prevalent in reflective writing with the consequence that 'real' learning does not occur and lacks '*practical synthesis'* (Watts, 2015, p. 368). In such cases, the reflective output does not meet the real-life needs for professional practice bringing the whole learning strategy into question. While there will always be learners who will take a strategic or superficial approach, it is important that reflective assessments do not serve as an 'educational iatrogenesis', where poor learning is caused by the same assessment that seeks to drive and measure it. To avoid this, reflective assessment design should consider the nature of target learners, the focus of the reflective assessment, and how this drives content that demonstrates wider learning *and* meets the intended learning outcomes. As previously discussed, resurrecting the 'reflective zombie' impairs the authenticity of the intended learning. Concerns regarding the sharing of authentic and real-life experiences are not new. Cotton's 'hegemonic discourse' (Cotton, 2001, p. 512) points to a potential for learners to feel 'policed' by reflective assessments and an educator would be well served to consider the extent to which they expect learners to share especially when analysing negative events in practice. Rolfe & Gardner (2006, p. 593) also describe a difference between learners who choose to adopt either an 'epistemological' or 'ontological' approach towards reflection. Pluralist voices espouse that learners have different approaches to both the context of a reflection, but also the components they wish to focus on. Some will choose to reflect on what they did and how they thought more than how they felt or how they were in practice. Forcing a focus on thoughts and feelings might not help the reflective process if the ways of practice are more important. The context of the problem or event might not require an analysis of feelings.

These limitations infer that the sharing of emotions can lead to poorer engagement in the reflective process and create unequal opportunities for some learners to shine. Worse still, emotional dissonance can occur, and this needs careful, individualized risk assessment to support students who may find themselves moving towards an emotional crisis (Watts, 2015). Educators are encouraged to consider which models and rubrics to use to guide student reflections and how much they involve the emotions

and sharing of feelings. For example, Gibb's Cycle (Gibbs, 1988) is often recommended in the development of reflective assessments but while this model comprises a focus on feelings, Driscoll's 'What?' model does not (Driscoll & Teh, 2001). It is important to consider which is the best fit in the context, given the focus of the reflective assessment and the learners involved.

Identification of risk associated to reflective practices remains an under researched area in educational (Mann et al., 2009). Whether a focus on sharing of and/or engagement in the affective components of practice should be contingent to effective reflection, is at least contentious. Educators should be highly critical of mandating emotional exploration in reflection and consider providing choice to the learner. This can be through a simple change to wording in rubrics that upholds the ideas of universal design for learning for all, as opposed to special measures for few. One example of such a small change is to provide choice to a learner exemplified by Monbec et al. (2021, p. 11), who in their rubric for student nurses, chose the words '*May* express feelings about or evaluation of the encounter' to describe a competent reflector. Whether reflective assessments are appropriate and inclusive of neurodiverse individuals represents a gap in the research and a novel question for further empirical exploration. Educators should avoid a blanket approach that associates lacking attendance to emotions and absence of sharing feelings with ineffective or poor reflection.

Formative vs. Summative Assessment

Sadler (2009) discussed several factors that might lead to a less positive learning and assessment strategy when employing analytic and holistic rubrics for grading divergent assessments (assessments for which there is no convergent single 'right' answer). These included concerns that applying rubrics was perceived as a more rigorous and robust method to judge performance in assessment tasks, when in fact grading is still likely to be inherently weak and lacking validity. Yet, an unfortunate and unintended problem with integrating rubrics for assessing reflective writing has arisen from misunderstandings of the intentions and legacy behind their development. Most developers of models, frameworks, and rubrics for reflection, did not intend them to be used for summative assessment (Ng et al., 2015). Contrastingly, they were intended to improve a learner's ability to reflect through a formative set of guiding principles. Anecdotally, it seems that at some point, this became misunderstood, with many formative rubrics and frameworks now being used for the summative assessment of reflective writing. Research has found the use of existing validated rubrics has only poor to moderate inter-rater reliability in the summative assessment of reflective writing (Soemantri et al., 2022), with the need for up to five assessors and 14 previous essays to reliably assess the reflective competence of one student (Moniz et al., 2015). The pragmatic educator must therefore assure the implementation of rubrics, models and frameworks is as intended, and understand the limitations of introducing fair and reliable marking models in high stakes reflective assessments. Currently, there is a lack of evidence to support the use of validated reflective rubrics to grade reflective assessment, and their developing authors, such as Wald et al. (2012), have been explicit in deterring their use for summative assessment purposes.

While authors such as Lucas et al. (2019) continue to work on the reliable quantification of scoring rubrics, educators are encouraged to subject reflective writing to formative or low stakes assessment for the time being. Formative assessment helps students understand their own experiences in practice (Smith & Trede, 2013). Formative assessment provides a parallel to the ethos of lifelong learning in practice and allows educators to avoid conflicting with pluralist voices who argue that, philosophically, reflection should not be subjected to quantifying measures (Ng et al., 2015). Formative assessment also

allows for the gradual development of reflective practice in a neurodiverse population of students. Beyond the potential inequity of sharing feelings in assessment is an added dimension around the assessment mode itself. Educators should be aware that not all academically able writers are reflective, and not all reflective learners are academically able writers (Sumsion & Fleet, 1996).

Problems arise in creating a pragmatic best practice approach to reflective writing assessments in higher education. Summative assessment is regularly used to motivate learners to meet defined learning outcomes for professional registration. Common educational practices include providing learners with previous years' summative assessment submissions as exemplars of 'good' reflections. For these to be valid and reliable guides to high stakes summative assessment, further work is needed to ensure that grading measures are also valid and reliable. To date, an empirical basis for such validity and reliability has yet to be realised. The heterogeneity of definitions and processes of reflection means a robust platform on which to judge a 'good' reflection does not exist. For this to occur needs well designed research, philosophical pragmatism, and a clear identification of defining terms. The literature contains a startling array of different assessment strategies in reflection and reflective writing throughout professional programmes. While this variety may improve student learning, this can also create confusing and variable implementation in higher education curricula, with staff and students being uncertain of the conventions of reflection across different subjects (Chan & Lee, 2021). Lacking robust evidence for reductionist approaches to assessing reflection juxtapose vividly with related counterarguments from pluralist voices, so that pragmatism might be seen as a reasonable middle ground in meeting the professional needs of students. However, those who speak with a pragmatic voice must be aware that the validity and reliability of summative assessment for reflective writing is beyond questionable. Until further research is completed, it is hard to use existing evidence to strongly support the use of validated rubrics for the assessment of reflective writing or indeed. Furthermore, high stakes summative assessment of reflective writing should be critically challenged.

Feedback to Develop Reflective Capability

A number of questions remain as to how students might understand and develop a high-quality written reflection, including its content and focus. An interesting distinction has arisen between reflection as an output and reflective practice as a process, and there is a risk in muddying the waters between process driven and content driven assessment. If providing guidance or a framework helps learners, then educators are more able to assess their ability to demonstrate reflective practice as a skill, as opposed to the content of the outputs from that process. Assessment that looks to assess reflection as a skill or competence, focuses on the steps required for effective reflection to occur, and as such is process driven. It is biased towards assessing the practice of reflection and as such risks the phenomenon of the 'reflective zombie'. Given the troubling lack of evidence for reflection as a science, the pragmatist assessor is encouraged to feedback more about how a learner learns from the practice of reflection than the oft idiosyncratic and personalised content of the reflection itself. For the latter, there is lacking evidence to support an assessor's value judgement (De La Croix & Veen, 2018). Moon's (2007, p. 194) 'depth dimension' appears as a comfortable middle ground. Depth has been applied to varying degrees in validated rubrics and appears as a platform to describe performance beyond process-based assessment, assessing the depth of learning without judging the content. Evaluating the appropriate depth of reflection is advocated here, as it is able to assess across diverse experiences. It also accounts for the fact that not all reflective learning experiences and contexts need to be subjected to the same depth for learning to occur. The inclusion

of depth as a criterion might appeal to pluralist and pragmatic voices which speak to assessments that 'describe how reflection takes place […] rather than prescribe what counts as 'good reflection'' (De La Croix & Veen, 2018, p. 398).

Pragmatic educators might like to consider facilitating reflection as praxis (a learner demonstrating individualised applications of reflective theory in practice) which leads to phronesis (sound judgement and evidence of practical wisdom). Neuroscientific and theoretical evidence supports specific resources and behaviours needed for effective reflection to occur, including providing clear guidance for students when being asked to complete reflective assessments, and encouraging time, space and putting in the effort of thinking upon thinking. The pragmatic voice should reconcile both reductionist and pluralist perspectives of reflection and advocate assessment briefs that describe the need for these resources and behaviours, and design rubrics that use language to describe how effective reflection looks (De La Croix & Veen, 2018). Assessment briefs and rubrics should describe a formative pathway for the process of reflection, upholding calls to avoid assessing content (Stewart & Richardson, 2000).

Impact on Developing Professional Practice

Despite its strong presence in higher education curricula, little evidence effectively connects the value of reflection and reflective assessment to the outcomes for service users, customers, and patients in the workplace (Mann et al., 2009; Uygur et al., 2019). While examples of past research exist that conduct a longitudinal assessment of educational interventions (Stupans et al., 2012), often these focus only on the educational outcomes, for example reflective writing, rather than more specific practice relevant outcomes that would emerge as applied actions from the learning from the reflection. Winkel et al. (2017) describe the importance of reflection towards developing empathy and communication skills, but empirical evidence to support a cause-and-effect relationship is lacking. Conversely, Watts (2015, p. 367) describes a potential for 'self-vertigo' evoked by critical reflections that transform a learner's core belief system and asks whether emotional dissonance following reflection on the self and negative events in practice, might lead to a desire to disconnect from the very profession to which learners seek to belong. This adds weight to an argument for more practice relevant descriptors in rubrics that intend to facilitate professional learning and belonging. For example, assessment rubrics might include reference to certain domains or action plans for professional practice (e.g. quality improvement, communication or leadership). Dunne et al. (2016) identified that practice-based written reflections often failed to link theory to practice and added a specific criterion to their rubric to nurture work-place relevance in their learners' practice based blogs.

The inclusion of reflection and reflective assessments in professional programmes, based on practice-based outcomes, is at least curious, if not questionable. Subsequently, more research is required to evaluate the impact of reflective practice on personal and service level outcomes (Winkel et al., 2017). In terms of meeting professional regulations, however, reflection is imperative, and despite ongoing debate, most authors agree that reflection is of value when embedded in a process of lifelong learning. Driscoll & Teh (2001, p. 98) state that reflective practice, 'Reminds qualified practitioners there is no end point to learning about their everyday practice.' Educationalists agree that learners need scaffolded educational interventions that bring them from a technical novice to a reflective expert practitioner, but not necessarily the best way for them to get there. With a reconciliation of both the reductionist and pluralist voices in the processes and outcomes of reflection, a pragmatic educator should be able to use the evidence here

to inform the design of rubrics that take learners from zombie to human, praxis towards phronesis and technical towards emancipatory professional practice.

Implications for Designing Rubrics for Reflective Writing

The persistence of a reductionist voice towards assessing reflection in higher education is unlikely to change any time soon, and a pragmatic voice invites educators to take evidence informed steps to apply best practices. These steps can ensure safety, inclusive learning, and fairness in reflective assessments. As little evidence exists to deter the use of rubrics in educational settings (Panadero & Jonsson, 2020), and recent reviews have suggested their use for reflective writing (Lim et al., 2023), more research is needed to support their ongoing use, and the design and evaluation of rubrics, both custom and generic, will only serve to increase the bank of supportive evidence that exists. Chan & Lee's (2021) recent literature review synthesised findings to provide four distinct needs for developing 'reflection literacy'. These needs are shared across different stakeholders in higher education including learners, teachers, and assessors. They consist of academic literacy, reflective knowledge and skills, assessment literacy and socio-emotional insights (Chan & Lee, 2021, p.12). The combination of these recommendations for facilitating reflection and reflective practices encapsulate a praxis of 'reflection literacy' and are integrated here to help frame a 'toolkit' for the development and implementation of student-facing assessment rubrics.

Explain Resources Needed Reflection to Occur

Ensure students understand the basic conceptual and neuroscientific processes of reflection. Explain that they will need the time, quiet space and energy required to engage in reflective practices. Last thing, on a Friday afternoon, might not be a good time. This will allow them to consider the development of lifelong praxis.

Keep It Low Stakes and Formative

Consider the appropriateness of formative and summative strategies. Be strongly skeptical of using validated instruments in high stakes graded reflective assessments. The evidence to support grading these is lacking.

Know Your Voice

Educators should consider their own voice and that of their institutional culture in the design of reflective assessment rubrics. They should select assessment language that clearly speaks to the voice of the assessment, identifying whether the learning outcomes resonate with a process-based, reductionist voice for reflection, a pluralist stance and content-based voice for reflection, or a more pragmatic reconciliation of the two.

Know Your Learners

Written reflection is only one mode of reflective assessment. Some learners will find this more challenging and may need more formative guidance and support. Decide whether your learners need to focus on the 'self' and thoughts and feelings early in their higher education careers to help mitigate undesirable side effects or hidden curriculums to the reflective assessments. While not covered in this chapter, alternative modes of formative assessment may facilitate student to produce other artifacts of reflective practice beyond reflective writing (eg creative inquiry such as art, poetry and story or talk-aloud). These can provide artefacts about which to write later. Multimodal assessment allows students who struggle with writing or expressing emotions to flourish. An artefact will show how a learner has perceived reflection to have been done or a praxis. The product of learning may be better described through them talking or writing about the artefact.

Choose Your Rubric

Educational theory would postulate that rubrics in the early years of higher education need to describe a procedural scaffold that describes a praxis of reflection. Therefore, learning the ropes of process-based reflective practice in early undergraduate years seems best suited to analytic rubric design. Higher education needs to encourage transformation in learners, allowing them to identify contextual paper-bags out of which they can punch their way in their own development. These may be better served by holistic rubrics where markers might make a more expert judgement on content-based outputs which demonstrate a practice phronesis. Formative assessment using tools such as the Reflection-on Action holistic rubric also saves time for educators (Miller-Kuhlmann et al., 2015) and skips the need to provide developmental feedback, which has hopefully been covered in earlier years of adult education.

Keep It Real

Challenge reflective assessment strategies that are either tenuously linked to real life practice or likely to drive a strategic approach to learning. Reflective assessment should mirror the challenges of professional contexts, learner diversity and meet meaningful learning needs for the learner and the end user of their practice.

Provide Clear Guidance and Choice

Many students, especially in early years of higher education, will want to use a framework. Provide one that aligns appropriately with the voice of your assessment and ensure it allows for exploration of relevant areas. For example, some frameworks and validated rubrics see thoughts and feelings as contingent on the process of reflection. When possible, provide choice around the sharing of feelings and emotions. If the task allows, provide students choice as to who they share their reflections with and whether they focus on the ways of being or ways of doing in practice. If you must use a certain framework for the purpose of meeting professional regulations, be explicit to learners about what is required. If the reflective writing needs to be supported by external evidence (eg literature and research or practice data) make this explicit in the assessment brief.

Be in it Together

Consider co-designing a formative rubric with your students. Careful and collaborative choice of the language of rubric design helps mitigate for some of the reflective contentions in current educational practice, particularly language that provides choice upon what to reflect and describes content upon which the learner is expected to focus and share.

Promote Peer Reflection

'Self' reflection or reflection are often conceptualised as a purely first-person activity, but as such it is inherently unreliable when based on one person's unchallenged and bias unilateral view of the world. Get learners to talk to each other about their experiences and provide advice and feedback. This will help mitigate self-deception or unhelpful rumination instead of useful reflection and action planning.

Encourage Depth and Holism

Perhaps most critically, educators should suspend judgement about what makes a good or bad reflection. The idiosyncrasies of individual reflections and outputs, the lack of research to support assessment of content, and the lack of evidence to support its impact on patient outcomes, means that educators might not be best placed to make such judgements. Instead, celebrate the depth of reflection and provide guidance on how to develop towards the next level. Choose rubric structure and language that describes how to reach reflective depth (novice learners) and appreciation of sociocultural practice and holism (advanced learners). Ng et al. (2015, p. 461) identify two 'theoretical orientations' for reflection in medical education - reflection on the epistemology of practice (the doing and how it is done) and reflection as a more critical social enquiry (looking at the wider influences on practice trends and conventions). Try to facilitate the learner's development from the former to the latter as they become more proficient. This moves a learner from the technical domain of doing the right things the right way, towards the emancipatory domain of questioning whether the right things are, in fact, the right things to be doing.

REFERENCES

Alsina, Á., Ayllón, S., & Colomer, J. (2019). Validating the Narrative Reflection Assessment Rubric (NARRA) for reflective narratives in higher education. *Assessment & Evaluation in Higher Education*, *44*(1), 155–168. doi:10.1080/02602938.2018.1486391

Alsina, Á., Ayllón, S., Colomer, J., Fernández-Peña, R., Fullana, J., Pallisera, M., Pérez-Burriel, M., & Serra, L. (2017). Improving and evaluating reflective narratives: A rubric for higher education students. *Teaching and Teacher Education*, *63*, 148–158. doi:10.1016/j.tate.2016.12.015

Aronson, L., Niehaus, B., Hill-Sakurai, L., Lai, C., & O'Sullivan, P. S. (2012). A comparison of two methods of teaching reflective ability in Year 3 medical students. *Medical Education*, *46*(8), 807–814. doi:10.1111/j.1365-2923.2012.04299.x PMID:22803758

Brookhart, S. M. (2013). *How to create and use rubrics for formative assessment and grading*. ASCD.

Chan, C. K. Y., & Lee, K. K. W. (2021). Reflection literacy: A multilevel perspective on the challenges of using reflections in higher education through a comprehensive literature review. *Educational Research Review*, *32*, 100376. doi:10.1016/j.edurev.2020.100376

Cheng, M. W. T., & Chan, C. K. Y. (2019). *An experimental test: Using rubrics for reflective writing to develop reflection*. Elsevier. doi:10.1016/j.stueduc.2019.04.001

Clouder, L., Karakus, M., Cinotti, A., Ferreyra, M. V., Fierros, G. A., & Rojo, P. (2020). Neurodiversity in higher education: A narrative synthesis. *Higher Education*, *80*(4), 757–778. doi:10.100710734-020-00513-6

Cotton, A. H. (2001). Private thoughts in public spheres: Issues in reflection and reflective practices in nursing. *Journal of Advanced Nursing*, *36*(4), 512–519. doi:10.1046/j.1365-2648.2001.02003.x PMID:11703545

Coulehan, J., & Granek, I. A. (2012). Commentary: "I Hope I'll continue to grow": Rubrics and reflective writing in medical education. *Academic Medicine*, *87*(1), 8–10. doi:10.1097/ACM.0b013e31823a98ba PMID:22201632

De La Croix, A., & Veen, M. (2018). EYE-OPENER The reflective zombie: Problematizing the conceptual framework of reflection in medical education. *Perspectives on Medical Education*, *7*(6), 394–400. doi:10.1007/S40037-018-0479-9 PMID:30353284

Devlin, M. J., Mutnick, A., Balmer, D., & Richards, B. F. (2010). Clerkship-based reflective writing: A rubric for feedback. *Medical Education*, *44*(11), 1143–1144. doi:10.1111/j.1365-2923.2010.03815.x PMID:20946509

Dewey, J. (2006). How we think. In *How we think*. D C Heath. doi:10.1037/10903-000

Donaghy, M. E., & Morss, K. (2009). *Guided reflection: A framework to facilitate and assess reflective practice within the discipline of physiotherapy*. Taylor and Francis. doi:10.1080/095939800307566

Driscoll, J., & Teh, B. (2001). The potential of reflective practice to develop individual orthopaedic nurse practitioners and their practice. *Journal of Orthopaedic Nursing*, *5*(2), 95–103. doi:10.1054/joon.2001.0150

Dunne, J., & Ryan, S. M. (2016). Enhancing Professional Development and Supporting Students on Work-Placement by Peer-Peer Learning Using an Online Reflective Blog Assessment. *Irish Journal of Academic Practice*, *5*(1), 1. doi:10.21427/D7HT51

General Medical Council. (2023). *Good medical practice*. GMC. www.gmc-uk.org/guidance

Gibbs, G. (1988). Learning by Doing : A Guide to Teaching and Learning Methods. *Further Education Unit*. [REMOVED HYPERLINK FIELD]

Greenfield, B., Bridges, P., Phillips, T., Adams, E., Bullock, D., Davis, K., Nelson, C., & Wood, B. (2015). Reflective Narratives by Physical Therapist Students on Their Early Clinical Experiences: A Deductive and Inductive Approach.[]. *Journal, Physical Therapy Education*, *29*(2), 21–31. https://journals.lww.com/jopte. doi:10.1097/00001416-201529020-00005

Grierson, L., Winemaker, S., Taniguchi, A., Howard, M., Marshall, D., & Zazulak, J. (2020). The reliability characteristics of the REFLECT rubric for assessing reflective capacity through expressive writing assignments: A replication study. *Perspectives on Medical Education*, *9*(5), 281–285. doi:10.1007/S40037-020-00611-2 PMID:32803530

Hatton, N., & Smith, D. (1995). Reflection in teacher education: Towards definition and implementation. *Teaching and Teacher Education*, *11*(1), 33–49. doi:10.1016/0742-051X(94)00012-U

Health and Care Professions Council. (2023). *Standards*. HCPC. https://www.hcpc-uk.org/standards/

Johns, C. & Freshwater, D. (1998). *Transforming nursing through reflective practice*. 222.

Kolb, D. A. (1984). *Experiential learning : experience as the source of learning and development*. Elsevier.

Lieberman, M. D. (2012). What zombies can't do: A social cognitive neuroscience approach to the irreducibility of reflective consciousness. In *Two Minds*. Dual Processes and Beyond. doi:10.1093/acprof:oso/9780199230167.003.0013

Lim, J. Y., Yew, S., Ong, K., Yan, C., Ng, H., Li, K., Chan, E., Yi, S., Wu, E. A., So, W. Z., Jin, G., Tey, C., Lam, Y. X., Lu, N., Gao, X., Lim, Y. X., Yong, R., Tay, K., Tze, I., & Krishna, R. (2023). A systematic scoping review of reflective writing in medical education. *BMC Medical Education*, *23*(12), 12. doi:10.118612909-022-03924-4 PMID:36624494

Lucas, C., Smith, L., Lonie, J. M., Hough, M., Rogers, K., & Mantzourani, E. (2019). *Can a reflective rubric be applied consistently with raters globally? A study across three countries*. Science Direct. doi:10.1016/j.cptl.2019.06.004

Maloney, S., Tai, J. H. M., Lo, K., Molloy, E., & Ilic, D. (2013). Honesty in critically reflective essays: An analysis of student practice. *Advances in Health Sciences Education : Theory and Practice*, *18*(4), 617–626. doi:10.100710459-012-9399-3 PMID:22926807

Mann, K., Gordon, J., & MacLeod, A. (2009). Reflection and reflective practice in health professions education: A systematic review. *Advances in Health Sciences Education : Theory and Practice*, *14*(4), 595–621. doi:10.100710459-007-9090-2 PMID:18034364

Miller-Kuhlmann, R., Osullivan, P. S., & Aronson, L. (2015). Essential steps in developing best practices to assess reflective skill: A comparison of two rubrics. Taylor & Francis. doi:10.3109/0142159X.2015.1034662

Monbec, L., Tilakaratna, N., Brooke, M., Siew, T., Lau, Y., Shih, C., & Wu, V. (2021). *Designing a rubric for reflection in nursing: a Legitimation Code Theory and systemic functional linguistics-informed framework Designing a rubric for reflection in nursing: a Legitimation Code*. Taylor & Francis. doi:10.1080/02602938.2020.1855414

Moniz, T., Arntfield, S., Miller, K., Lingard, L., Watling, C., & Regehr, G. (2015). Considerations in the use of reflective writing for student assessment: Issues of reliability and validity. *Medical Education*, *49*(9), 901–908. doi:10.1111/medu.12771 PMID:26296406

Moon, J. (2007). Getting the measure of reflection: Considering matters of definition and depth. *Journal of Radiotherapy in Practice*, *6*(4), 191–200. doi:10.1017/S1460396907006188

Ng, S. L., Kinsella, E. A., Friesen, F., & Hodges, B. (2015). Reclaiming a theoretical orientation to reflection in medical education research: A critical narrative review. *Medical Education*, *49*(5), 461–475. doi:10.1111/medu.12680 PMID:25924122

Nguyen, Q. D., Fernandez, N., Karsenti, T., & Charlin, B. (2014). What is reflection? A conceptual analysis of major definitions and a proposal of a five-component model. *Medical Education*, *48*(12), 1176–1189. doi:10.1111/medu.12583 PMID:25413911

Panadero, E., & Jonsson, A. (2020). A critical review of the arguments against the use of rubrics. In *Educational Research Review* (Vol. 30). Elsevier Ltd., doi:10.1016/j.edurev.2020.100329

Rogers, J., Peecksen, S., Douglas, M., & Simmons, M. (2019). *Reflective Practice International and Multidisciplinary Perspectives Validation of a reflection rubric for higher education*. Taylor & Francis. doi:10.1080/14623943.2019.1676712

Rogers, R. R. (2001). Reflection in Higher Education: A Concept Analysis. *Innovative Higher Education*, *26*(1), 37–57. doi:10.1023/A:1010986404527

Rolfe, G., & Gardner, L. (2006). "Do not ask who I am...": Confession emancipation and (self)-management through reflection. In Journal of Nursing Management, 14 (8), 593–600. https://doi.org/doi:10.1111/j.1365-2934.2006.00717.x

Sadler, D. R. (2009). Indeterminacy in the use of preset criteria for assessment and grading. *Assessment & Evaluation in Higher Education*, *34*(2), 159–179. doi:10.1080/02602930801956059

Schön, D. A. (1983). *The reflective practitioner : how professionals think in action*. Taylor & Francis.

Smith, M., & Trede, F. (2013). *Higher Education Research & Development Reflective practice in the transition phase from university student to novice graduate: implications for teaching reflective practice*. Taylor & Francis Online. doi:10.1080/07294360.2012.709226

Soemantri, D., Mustika, R., & Greviana, N. (2022). Inter-Rater Reliability of Reflective-Writing Assessment in an Undergraduate Professionalism Course in Medical Education. *Education in Medicine Journal, 14*(1), 87–97.

Stupans, I., March, G., & Owen, S. M. (2012). *Assessment & Evaluation in Higher Education Enhancing learning in clinical placements: reflective practice, self-assessment, rubrics and scaffolding*. Taylor & Francis Online. doi:10.1080/02602938.2012.658017

Sumsion, J., & Fleet, A. (1996). Reflection: Can we assess it? Should we assess it? *Assessment & Evaluation in Higher Education*, *21*(2), 121–130. doi:10.1080/0260293960210202

Torrance, H. (2007). Assessment as learning? How the use of explicit learning objectives, assessment criteria and feedback in post-secondary education and training can come to dominate learning. *Assessment in Education: Principles, Policy & Practice*, *14*(3), 281–294. doi:10.1080/09695940701591867

Uygur, J., Stuart, E., De Paor, M., Wallace, E., Duffy, S., O'Shea, M., Smith, S., & Pawlikowska, T. (2019). A Best Evidence in Medical Education systematic review to determine the most effective teaching methods that develop reflection in medical students: BEME Guide No. 51. *Medical Teacher*, *41*(1), 3–16. doi:10.1080/0142159X.2018.1505037 PMID:30634872

Wald, H. S. (2015). Refining a definition of reflection for the being as well as doing the work of a physician. *Medical Teacher*, *37*(7), 696–699. doi:10.3109/0142159X.2015.1029897 PMID:25897706

Wald, H. S., Borkan, J. M., Taylor, J. S., Anthony, D., & Reis, S. P. (2012). Fostering and evaluating reflective capacity in medical education: Developing the REFLECT rubric for assessing reflective writing. *Academic Medicine*, *87*(1), 41–50. doi:10.1097/ACM.0b013e31823b55fa PMID:22104060

Watts, L. (2015). An Autoethnographic Exploration of Learning and Teaching Reflective Practice. *Social Work Education*, *34*(4), 363–376. doi:10.1080/02615479.2015.1016903

Winkel, A. F., Yingling, S., Jones, A.-A., & Nicholson, J. (2017). Reflection as a Learning Tool in Graduate Medical Education: A Systematic Review. *Journal of Graduate Medical Education*, *9*(4), 430–439. doi:10.4300/JGME-D-16-00500.1 PMID:28824754

Chapter 16

Pedagogical Potential and Didactic Limitations of Assessment Rubrics:
An Example From Medical Education

Murat Tekin

iD https://orcid.org/0000-0001-6841-3045

Çanakkale Onsekiz Mart University, Turkey

ABSTRACT

High-level cognitive skills are often demonstrated at the performance level. For this reason, performance assessment has become an important element of educational assessment. One of the biggest problems in open-ended questions, oral exams, and performance evaluation is to develop a scoring method that will ensure consistency between raters. At this point, the rubric emerges as a functional scoring tool. Rubrics have many known advantages such as defining the elements and qualities of performance that should be exhibited in educational assessment, consistency between raters, and supporting teaching. On the other hand, there are situations that create barriers in use, such as difficulty in preparation, time consuming to prepare, requiring expertise, and defining the qualifications by distributing them in a balanced way. However, it cannot be said that these are the only obstacles in the use of rubrics. At the same time, educator typology in lesson process and trainer profiles, educational beliefs, and educator roles can be counted among the important barriers in using rubric.

INTRODUCTION: RUBRIC

When a performance assessment is made, evaluators are faced with a structured outcome or responses created by the learner instead of responses of students chosen from among the options. Although it is more complex than scoring student answers selected from these options as in multiple-choice exams, the outcome or structured answers that emerge as a result of performance must be scored. In scoring a created outcome or response, the evaluation criteria used in demonstrating the adequacy of this outcome

or created response gain significant importance. Therefore, criteria take the focus in performance evaluation. The scoring procedures used to evaluate students' responses to performance tasks are generally referred to as rubrics (Popham, 2017).

Rubric definitions have been made in various ways. According to Haladyna (1997), who defines observations as a tool that allows scoring by categorizing, only experts should do the scoring. Rubrics should be used to evaluate more abstract concepts such as subject integrity and quality rather than concrete behaviors such as size and length that anyone can easily do. Moscal (2000) made another definition of rubrics and stated that it is a scoring key developed by the trainer or evaluator to analyze the students' and their outcomes. They can evaluate the qualifications expected from the learner in a broad framework and present a judgment and a frame of reference. Miller, Linn, and Gronlund (2009) defined rubrics as a measurement tool that clearly determines whether a performance is accepted by students and educators.

A rubric is a set of clear expectations or criteria to help instructors and students focus on what is valuable in a topic or activity. Rubrics are often similar to checklists in that they list multiple criteria for a performance. However, unlike a checklist that simply lists the criteria, a rubric describes the expected performance level for each criterion. Rubrics set benchmarks for different performance levels, often descriptive rather than numerical. Descriptions help educators focus their teaching and student work on key aspects of the rubric. Explanations also help students better understand what trainers expect from them for a particular performance or outcome (Russell & Airasian, 2012).

In line with all these definitions, it can be said that the rubric;

- Determines the performance-related situation
- Determines if the performance is acceptable or not
- Is a reference
- Functions as a key for scoring
- Is a set of criteria that can grade performance from strong to weak/good to bad
- Can make the expectations of the trainer very clear
- Provides more informative feedback about the learners' strengths and aspects that need improvement
- Supports learning, the development of skills, understanding, and good thinking

A rubric used to score students' responses to a performance assessment has at least three important features (Popham, 2017):

- Evaluative Criteria: Factors used to determine the quality of the learner's response.
- Descriptions of Qualitative Differences for all Evaluative Criteria: An explanation for making qualitative distinctions in the learner's response for each assessment criterion included in the rubric.
- Determining Which Scoring (holistic or analytical scoring) Approach to Use: In rubrics, the evaluation criteria can be in the form of holistic or criterion-based analytical scoring. The type of scoring to be used should be specified.

There are two types of rubrics: analytical and holistic. Analytical rubrics show important dimensions, features, and elements of learners' responses, enabling them to be scored. Holistic rubrics aim to evaluate the quality of learners' responses in general (Nitko & Brookhart, 2014). Based on the decision, a holistic rubric can be used to make broad decisions about student performances or an analytical rubric can be

used to focus on specific aspects of performance. The holistic rubric contains a single explanation for each performance level. This single statement usually focuses on the extent to which multiple criteria are met in performance. On the other hand, analytical rubrics include a separate explanation for each performance criterion and provide a separate score for each (Russell & Airasian, 2012).

Rubric Sections

A rubric generally consists of three sections: Assessment criteria, attribute definitions, and scoring strategies.

<u>Assessment Criteria:</u> Is used to distinguish whether the student's answers are acceptable or not. The rubric should be prepared according to the content of the skill to be evaluated. For example, if we are evaluating the learner's problem-solving skills, the evaluation criteria should be different than if we are evaluating the communication with the patient. Skill requirements determine the assessment criteria.

<u>Attribute Definitions:</u> Refers to identifying qualitative differences in the assessment of responses. The learner with the highest score in the assessment should not have made any mistakes in their answers.

<u>Strategies of Scoring:</u> Defines how scoring should be done. Rubrics can be created holistic or analytically.

Developing the Rubric

There are different suggestions in different sources for developing a rubric. According to Haladyna (1997), the stages of rubric development are as follows:

1) Determining the purpose of the rubric: This purpose can be not only to determine the performance but also to detect student mistakes and give feedback.
2) Clearly specifying the criteria to be scored: Students need to know according to which criteria they will be evaluated.
3) Deciding which type of rubric to use: Holistic or analytical; generic or task-specific.
4) Making a draft of the rubric.
5) Making corrections.
6) Application for trial purposes.
7) Evaluation of results related to student performance using a rubric.
8) Determining the consistency and reliability of the scoring.
9) Making corrections for future uses.

According to Moskal (2000), the following should be determined to develop a rubric.

1) Clearly defining the adequate or highest level in students' work, assignments, presentations, or desired performance; or defining the qualities of skillful performance.
2) Scoring criteria at different levels of the defined attributes should be specified.
3) Deciding what type of rubric will be used for these attributes we have defined: Holistic or analytical?
4) If an analytical rubric is to be developed, defining and developing the score levels of the elements of the trait under consideration. If a holistic rubric is to be developed, examine the levels to be combined, create each level, and ultimately devise a defined score.

Although the ways of developing a rubric have been expressed in different ways with different perspectives, it is possible to see that there are common points between them. The priority is to determine the purpose. In line with the objective, it is necessary to determine the qualifications and/or sub-qualifications, that is, adequately define the elements of the expected skill. It is necessary to decide on their grading level by grading the defined elements as upper, middle, and lower qualities. Then comes the appropriate rubric type: holistic, analytical, generic, task-specific, or mixed scoring type, should be determined. When the rubric is ready, it is necessary to share it with the learners, evaluate it using established criteria, and give feedback to the learner after the application.

The rubric must be clear and understandable. Expected learner performances should be clearly indicated. Statements should be prepared carefully, without errors or areas of misunderstanding, and attention should be given to be free from prejudice.

Benefits of Using Rubrics

Rubrics are scales that ensure the consistency of the scoring performed and indicate the success standards of the students in a given task. Rubrics attempt to answer the following questions (Nitko & Brookhart, 2014): What learning targets and criteria should be evaluated? What should be the achievement levels of these learning targets and criteria? What type of rubric should be used (analytical or holistic)? Should students participate in the process of assessing their own performance? How should scoring be organized so that it takes less time and is effective? What needs to be recorded as evidence of assessment?

One of the main limitations of tests for performance assessment is the required scoring time. It takes time and effort for the student to perform the tasks assigned within the scope of performance assessment. Similarly, scoring takes time and effort for raters. The scoring of true-false, multiple-choice, and short-answer questions is easy, objective, and fast. However, evaluation of projects, portfolios, or performances forces trainers to make difficult choices in converting the learner's effort, participation, and cooperation into points. Relying on easy or less controversial scoring to evaluate performance can turn a well-thought-out and original performance measure into an incorrect assessment option. Your goal when scoring performance tests is to justify the time spent developing them and the effort students put into completing performance tasks. You can achieve this goal by developing carefully crafted scoring systems called rubrics. Rubrics are scoring keys that, by careful evaluation, keep students at high achievement standards while minimizing the arbitrariness of your judgments (Kubiszyn & Borich, 2013).

The rubric helps develop a shared understanding of what is valuable in a performance by providing a description and sometimes examples of each performance level for each performance criterion. This common understanding increases the reliability of scores given by multiple raters by focusing raters on the same performance elements. Additionally, this shared understanding helps increase the validity of performance appraisals by helping students identify aspects of performance. Rubrics help educators set criteria to focus on what is important in the learning-teaching process, set criteria to focus on student assessments, increase consistency of assessment, reduce subjectivity in setting criteria and scoring, and ensure students understand the expected performance. Rubrics help students clarify instructors' expectations for performance, state what is important in a process or outcome, monitor and critique their own work, inform by detailing expected performance, and provide clearer performance information than the traditional "letter grading" system (Russell & Airasian, 2012).

Rubric Usage Difficulties and Disadvantages

As stated in the literature, rubrics are very useful and recommended tools for scoring the outcomes that emerge in performance evaluation. However, despite its known benefits, there are also disadvantages and some difficulties in using rubrics. These are;

- Rubrics take time to prepare
- Rubrics require expertise and experience in its preparation
- The raters may experience hesitations between points when rating the quality of the outcome
- Accordingly adjusting the degrees showing the quality of the outcome in a balanced way
- Identification of the components that show the quality of the outcome by thorough consideration

Tenam-Zemach and Flynn (2015) listed some of the disadvantages of rubrics as follows:

- Rubrics can limit evaluators' perceptions and prevent them from seeing and commenting on important aspects of the education.
- Some educators dislike such precise, prepackaged assessment tools.
- Trainers can compare their ratings with their peers. Thus, they can turn the evaluation process into a numbers-based competition to gain an advantage over their peers.
- The management may request evidence from trainers for each rating, making using rubrics time-consuming.

These difficulties or disadvantages can create obstacles or hesitations about using rubrics.

Barriers to the Use of Rubrics

Educator Typology in Lesson Process and Trainer Profiles

One of the most important concepts/phenomena brought by educational psychology to educational science is "individual differences" and the importance of these differences. Individual differences in education are a phenomenon that is generally taken into account by learners. However, educators have individual differences just as learners. Brown, Bakhtar, and Youngman (1984) divided trainers into 5 typologies in terms of course presentations. The first typology is called Oral Lecturer. This typology neatly structures their lessons but is not very keen on setting a set of written objectives for the lesson or telling their students the lesson's objectives. These trainers are confident in their verbal abilities and mainly communicate orally in their lessons. The second typology is the Exemplary Lecturer. This typology course is against reading and giving students too detailed information. They appear confident and organize their lessons according to a set of goals. In the presentation, they present the lesson in a logical framework. They use visuals to show the process and procedures. They reiterate, emphasize, and summarize when necessary. The third typology is Information Providers. They act as a newscaster in these typology classes and tend to read from their well-structured notes. Although they have the skills to use visual tools and are self-confident, their use of visuals in their presentations is weak. Instead, they give a lot of details on the topics and try to present them all. The fourth typology is the trainers that do not fit a certain pattern (Amorphous Lecturers). This typology has no set purposes. As a result, their

presentation styles are very weak compared to other typologies. The fifth typology is the Self-Doubters typology. They have difficulty selecting and structuring materials. They do not stay close to the content and structure of their notes, and many do not feel that they have achieved the objectives by the end.

Educator profiles have been studied as well as educator typologies in terms of course presentation. Kremer and Hofman (1979) classified trainer profiles as follows: open-minded educator, internally directed educator, affectively oriented educator, cognitively oriented educator, externally directed educator, and close-minded educator. Educator profiles were also studied on integrating technology with education. Mama and Hennessy (2013) defined four educator profiles in using technology in education. These are Integrational, Incremental, Incidental, and Inimical profiles. The integrational profile is one that incorporates technology into classroom practices and tries to integrate it with pedagogy. The incremental profile is a profile that enriches existing classroom practices with the use of technology. The incidental profile is used occasionally and integrates in-class applications and technology. On the other hand, an inimical profile consciously stays away from technology in classroom pedagogical practices. There are different teacher typologies as well as different teacher beliefs.

Educator Beliefs

The nature of beliefs and their psychological bases have been studied for a long time. Beliefs are structured within a system. This system should be considered layered. While some beliefs are more external and superficial, others may be deeper. Deep beliefs are resistant to change, while superficial beliefs are more open to change. Rokeach (1968) thought of this system in beliefs as the nucleus of the atom and electron orbits (rings) outside it. He has foreseen 5 rings (5 types) from the center outwards. At the center is Type A Belief. This type of belief consists of basic premises about the physical and social environment and beliefs about the self. These beliefs are resistant to change. Type B Belief reflects one's self-image of one's own beliefs. Type C Belief is a sign of belief attributed to authoritarian figures such as family, friends, and politicians. Type D Beliefs are beliefs derived from authoritarian figures, such as political and religious, rather than from direct interaction with the object of the belief. Type E Beliefs are beliefs related to personal pleasure and satisfaction arising from the direct interaction between the individual and the belief object.

It can be said that there is a consensus on examining belief as a system within a cognitive structure (Fives & Buehls, 2012; Nespor, 1987; Rokeach, 1968). The system approach to beliefs is also valid in education when dealing with "educators' beliefs." Fives and Buehls (2012) stated that beliefs affect teachers' practices as a system. Therefore, belief as a psychological structure is very important in the education of trainers. Types of educational beliefs that affect classroom practices are context and environment, content and information, special teaching practices, teaching approaches, and beliefs about students. In studies, the structure and types of beliefs have been the focus of attention in terms of their reflection on the practices of the educators.

Törner (2002) named the trainer's disciplinary beliefs, global beliefs, beliefs about any field within the discipline, domain-specific beliefs, and beliefs about a subject within that field as subject matter beliefs. Examples of general beliefs are medicine-specific beliefs that answer questions such as "What is medicine, what is the nature of medical science, and how should medicine be taught." Examples of domain-specific beliefs can be given in sub-disciplines of medicine, such as "How should internal and surgical sciences be learned and taught." The answers to the questions of "How should muscles, nerves, and systems be taught," which are the course contents of anatomy, can be given as an example of subject

field beliefs. These three types of beliefs interact. For example, in terms of general medical education beliefs, a trainer who believes that "Medical knowledge is formed with experience and many practices" may reflect this belief to field-specific beliefs such as "Medicine should be taught by doing many practices at the patient's bedside." According to Kuntze (2012), since teachers' professional knowledge is structured episodically and each teacher's experiences, lives, and memories are separate, it is necessary to explain teaching with a belief system model. Teaching independent of belief systems may lead to incomplete or limited teaching evaluation.

Educator Roles

Education has undergone a significant change in its historical process. While the influence of behavioral theories in learning was strong in the past, the influence of cognitive-behavioral, cognitive, and humanistic learning theories has increased over time. While the traditional discipline understanding based on control and punishment used to be prevalent in classroom management, the impact of humanist and democratic class management approaches that put people in the center increased over time. While it was important to make all individuals the same in education in the past, it has gained importance to reveal individual differences as much as possible over time. This transformation in education has also led to a transformation in teacher roles.

Brophy (1985) suggested that educators see themselves primarily as instructors or socializers in their interactions with students and that educators' perceptions of this role affect how they interact with students. While educators in the role of teacher tend to react more negatively towards students who are inadequate, reluctant, or destructive in the teaching process, educators in the role of socializer tend to behave more negatively towards students who they see as hostile and aggressive or push them away as the educator tries to establish a relationship. Wubbels and Levy (1993) grouped the relationships between educators and students under two dimensions. The first is obedience to domination, and the other is opposition to cooperation. The ideal role in this classification is moderate rather than excessive dominance and moderate rather than excessive cooperation. Eight types of trainer behaviors are included under these two dimensions, which are represented as two axes. These educational behaviors are as follows: leader, helpful/friendly, understanding, giving freedom and responsibility to students, uncertain, dissatisfied, giving advice, and harsh. A leader educator notices what is going on, leads, organizes, and gives instructions. A helpful/friendly educator is someone educational, helpful, interested, and friendly. An understanding educator is a person who listens with interest, gives confidence, shows understanding, and is patient. An educator that gives freedom and responsibility to the students provides opportunities for independent study, gives freedom, and responsibility. The uncertain educator looks from where the wind will blow. Dissatisfied educators wait for silence, consider the pros and cons, and show dissatisfaction. Advice givers get angry, correct mistakes, and impose prohibitions. Finally, tough educators set rules and norms, act tough, and maintain silence in the classroom.

There may be hesitations, barriers, and inherent disadvantages and difficulties of rubrics as a scoring method in evaluation. However, hesitations and barriers to using rubrics may also be the educators' typologies, profiles, preferred teaching, assessment and evaluation types, their beliefs as educators, and their educational roles.

Medical Education as an Example

Rubric Usage Areas in Medical Education

It is known that there are some components of medical education. The first is the great accumulation of knowledge it has produced in the historical process. It has an ancient body of knowledge. This information should be transferred to every new physician who enters the profession. However, one of the debates at this point is how and what proportion of the dizzyingly updated body of information should be taught to physicians. The second element is the skills to be acquired. It is necessary to be able to examine the patient's body, to perform some interventional procedures (such as doing injections and intervening in a patient with cardiac), and even if the surgical branches are specialized, the necessary surgeries should be performed successfully. The third element of medical education can be called professionalism. Perceptions about their professions, their approach to patients and their work while practicing are accepted within professionalism. The learning that takes place under these components is expected to turn into the ultimate competence. In summary, medical education aims to develop certain competencies in the physician as professional staff. For this reason, most medical education falls within the scope of performance evaluation.

Medical education has developed greatly in the historical progress. Until the beginning of the 1900s, there was an education style that was more suitable for the master-apprentice order based on memorizing the information. Voices against this form of education have risen over time. As reflected in the report of Flexner (1910), demands such as increasing teaching and learning processes that strengthen integration, encouraging research and development, and supporting individual learning and making more standard assessments, and supporting professional identity have increased. These demands have evolved medical education towards an integrated education. Harden, Sowden, and Dunn (1984) formulated how to integrate medical education into the curriculum. According to them, integration is "the arrangement of what is taught in different academic courses and/or departments to relate or integrate them." Integration allows learners to question "what" and "how" they will learn. McMaster School of Medicine in Canada was the first faculty to replace the traditional education model with problem-based learning (PBL) in 1969 and apply it to a progressive and transdisciplinary integration program. Today, some medical faculties start vertical integration with clinical education of upper classes and horizontal integration starting from the first year. While some other faculties give more importance to horizontal integration in the 1st, 2nd, and 3rd grades, they structure medical education by constructing horizontal and vertical integration in the 4th and 5th grades.

In line with what has been said so far, education is carried out on the principle that medical education takes place based on performance and that the knowledge and skills of all departments within the scope of medical science are integrated and brought to the learners. Many educational elements in medical education are conducted based on case-based discussions, clinical reasoning, panels, bedside training, etc. Similarly, clinical reasoning problems, problem-based evaluations, case discussions, report review and evaluation, structured oral and written exams, internship reports that serve as a kind of portfolio, and many similar performance evaluation methods are used in the assessment and evaluation phase. Examples of these assessment methods and their rubrics are as follows:

Example One

Explanation: The following example has been prepared for students who are at the level of Intern Physician studying at the Faculty of Medicine. The example includes a performance task to perform and an analytical rubric to be used in scoring this task.

Performance Task

You work as a family physician at the Family Health Center No. 8 in the Central District of Çanakkale. On Monday, a 63-year-old retired male patient came to you. The patient stated that he has just moved to your service area and wants to register in your system. Therefore, this patient will be added to your list and will be under your follow-up. The patient said that he had high blood pressure and that his shoes had been tight recently due to swelling in his feet. Explain what you will do with this patient who applied to you. In your explanation, it is recommended to pay attention to the following points according to the principles of Periodic Health Inspection:

- Background of the patient
- Family history
- The current situation

Table 1. Analytic rubric

Medical Skill Components	Should be Improved (1)	Good (2)	Excellent (3)
Welcoming and communicating with the patient	Welcomed the patient, did not ask open-ended questions, asked only confirmation questions for which she/he would receive yes and no answers. Did not ask additional questions to obtain sufficient information. Could not receive and give feedback.	Welcomed the patient, indicated where to sit. She/He made eye contact. It was insufficient to ask open-ended questions. In communication with the patient, she/he interrupted the patient from time to time. When the patient did not provide sufficient information in her/his speech, she/he asked additional questions to obtain sufficient information. Couldn't get enough feedback and couldn't give it.	Greeted the patient with a smiling face, showed her/him a comfortable place to sit. She/He made eye contact. It facilitated communication with open-ended questions. She/he listened to the patient without interrupting. When the patient did not provide sufficient information in her/his speech, she/he asked additional questions to obtain sufficient information. Provided feedback showing understanding of the patient. Received feedback from patient confirming patient's understanding of their own explanations.
Ability to take anamnesis	While questioning the patient's complaints, she/he questioned them superficially without going into details. She/he did not question the background of patient. She/he did not question the patient's family history.	She/he questioned the patient's complaints. She/he did not question the background of patient. She/he did not question the patient's family history. She/he learned about the drugs the patient was using.	She/he made a detailed questioning of the patient's complaints. She/he did not question the background of patient. She/he did not question the patient's family history. She/he questioned in detail the drugs used by the patient, drug side effects, and drug use patterns. She/he made a general system inquiry of the patient. She/he determined the risks by making the necessary inquiries within the scope of the periodic health examination.
Physical examination skills	She/he measured the patient's arterial blood pressure without making the necessary preparations or did not measure at all. Did not perform cardiovascular system examination. She/he did not inspect all systems.	She/he measured the patient's arterial blood pressure without making the necessary preliminary preparation. Cardiovascular system examination was done. She/he did not perform a detailed examination of all systems.	She/he measured the patient's arterial blood pressure by making the necessary preliminary preparations. She/he made detailed cardiovascular system examination (checking the heart and arteriovenous system by inspection, palpation, percussion, auscultation). She/he made a detailed inspection of all systems.
Ability to identify investigations to be performed	She/he did not request tests from the patient or the tests they requested were missing for arterial hypertension and periodic health examination.	She/he requested testing for arterial hypertension. She/he requested a general examination for periodic health examination.	She/he requested the correct tests to be done for the patient with arterial hypertension. She/he determined the risk status of the patient and asked for the necessary tests within the scope of the periodic health examination.

Example Two

Explanation: The example below has been prepared for students studying in the Year 4 "Internal Medi-

cine Internship" at the Faculty of Medicine. The example includes an open-ended question and a holistic rubric to be used in scoring this question.

Open-Ended Question

Explain the principles to be considered in abdominal examination.

Table 2. Holistic rubric

4-Excellent • She/he explained the importance of inspection and explained the findings and features to be determined by inspection. • She/he explained why auscultation examination should be done before palpation. She/he explained the findings to be determined by auscultation and their importance. • She/he explained the importance of palpation. She/he explained the findings that can be detected by palpation. • She/he explained the importance of percussion examination. She/he explained the findings and features to be detected by percussion. • She/he explained how the examinations of the abdominal organs were performed.
3-Good • She/he explained the importance of inspection and explained the findings to be determined by inspection. • She/he explained why auscultation examination should be done before palpation. She/he explained the findings to be detected by auscultation. • She/he explained the findings that can be detected by palpation. • She/he explained the importance of percussion examination. She/he explained the findings to be detected by percussion.
2-Should be Improved • She/he explained the importance of inspection. • She/he explained the importance of auscultation. • She/he explained the importance of palpation. • She/he explained the importance of percussion.
1-Insufficient • She/he made definitions of inspection, auscultation, palpation and percussion. However, the definitions were confused and inaccurate.

Example Three

Explanation: The following example has been prepared for students studying in the Faculty of Medicine Year 4 "Family Medicine Internship". The example includes an open-ended question and a holistic rubric to be used in scoring this question.

Open-Ended Question

Explain the principles to be considered in communication during the incoming patient interview.

Table 3. Holistic rubric

4-Excellent • She/he explained the importance of welcoming the patient with a smiling face and placing her/him in a comfortable place (It provides the first positive prejudices and a sense of trust to be formed between the patient and the physician). • She/he explained the importance of making eye contact with the patient. (Confidence increases in the patient who feels that she/he is being listened to, and the patient's non-verbal expressions are captured). • Explained the importance of facilitating communication with open-ended questions (It allows the patient to express herself/himself comfortably). • She/he explained the importance of listening to the patient without interrupting (The patient understands that she/he is being listened to, allows her to express herself/himself comfortably) • The patient explained the importance of summarizing and giving feedback in the physician interview (Making the patient feel that he/she understands provides an opportunity to correct the information that the patient did not explain or explained incompletely and prevents misunderstanding). • The feedback explained the importance of understanding (Making sure that the topics discussed are understood by the patient).
3-Good • It is not sufficient to explain how the patient will be greeted and body language. • She/he explained the importance of facilitating communication with open-ended questions. • The patient explained the importance of summarizing and giving feedback in the physician interview. • Explained the importance of understanding feedback.
2-Should be Improved • In facilitating communication, open-ended questions, summarizing, giving feedback and getting feedback are not clear.
1-Insufficient • She/he just gave the definition of communication.

Medical Educator Teaching Trends and Possible Resistances to the Use of Rubrics

Faculty members define themselves as members of an academic discipline rather than educators (Kember, 1997). Medical faculty members should not be considered outside of this perception. Physicians are trained to provide health care. If they work as a lecturer in a university hospital, the duty of doing research and giving lectures is added to the duty of providing health services. However, physicians' extrinsic motivation for teaching is often low. The use of information technology in education and the application of student-centered strategies such as problem-based learning in medical education did not always coincide with the conceptualization and beliefs of the educators and even met with resistance (Vatansever, 2011).

Vatansever (2011) conducted a qualitative research with clinician faculty members in the medical faculty and determined that clinician faculty members mostly used didactic methods, questioning, and example-giving techniques in lessons. To teach practice-based learning objectives, bedside training associated with objectives are more frequently preferred. In addition, there are methods that physician faculty members do not have the opportunity to use, such as skill training with models, and watching students throughout the education. The metaphors of faculty members about learning and teaching reflect a student-centered perspective. Many of the metaphors show features of guidance, protection, reflection, and role modeling. The traditional transfer metaphor was expressed by a faculty member. Similar features to these metaphors of the participants can also be seen in the definitions of ideal trainers. It was observed that faculty members value being role models and are satisfied with their contribution to teaching. Participants also value innovative approaches not only in pre-graduate (the name given to the six years of medical education, undergraduate education), post-graduate education (the name given to the specialist training of medical education), and in-service training of non-university health profession-

als. It is also necessary to evaluate whether these orientations of clinical educators are reflected in their teaching strategies. The practice of didactic methods by faculty members expressing a student-centered orientation suggests an inconsistency between the orientations of clinical educators and their teaching strategies. Another inconsistency was identified in the teaching focus. Some faculty describing a student-centered orientation reported their teaching priorities and focus as subject-centered or fully aligned with the curriculum, neither of which are student-centered approaches. The main obstacles to an effective student-centered education were defined as workload, insufficient patient diversity in clinics, an excess number of students, and insufficient teaching skills.

Research has also shown that even in areas where performance-based education is at the highest level, the use of student-centered teaching methods and performance evaluations are in the background. Physician faculty members are more focused on teaching the subjects. This makes education more rote learning. Faculty members who use teaching principles, learner-centered teaching strategies, and performance evaluation also tend to prepare detailed rubric-like scoring keys. The long and tedious preparation process of rubrics and the difficulty of making definitions by considering the quality of the performance is overwhelmed by the workload of the physician faculty members and the priority they give to providing health services. Of course, the only reason is not the priority and intensity of the faculty members in providing health services. Perhaps this is the apparent reason. Actually, possibly the biggest obstacles to using performance evaluation and evaluating it with rubrics are the features described in this chapter's *"Educator Typology in Lesson Process and Educator Profiles", "Educator Beliefs," and "Educator Roles."*

REFERENCES

Akar, H., & Yildirim, A. (2009). Change in teacher candidates' metaphorical images about classroom management in a social constructivist learning environment. *Teaching in Higher Education*, *14*(4), 401–415. doi:10.1080/13562510903050152

Brophy, J. (1985). Teachers' expectations, motives, and goals for working with problem students. In C. Ames & R. Ames (Eds.), Research on motivation in education: Vol. 2. The classroom milieu (pp. 175–213). New York: Academic Press.

Brown, G. A., Bakhtar, M., & Youngman, M. B. (1984). Toward a typology of lecturing styles. *The British Journal of Educational Psychology*, *54*(1), 93–100. doi:10.1111/j.2044-8279.1984.tb00848.x

Fives, H., & Buehl, M. M. (2012). Spring cleaning for the "messy" construct of teachers' beliefs: What are they? Which have been examined? What can they tell us. *Apa Educational Psychology Handbook*, *2*, 471–499. doi:10.1037/13274-019

Flexner, A. (1910). Medical education in The United States and Canada a report to the Carnegie Foundation for the advancement of teaching. Bulletin Number Four. Boston: The Merrymount Press (Reproduces in 1972).

Haladyna, T. M. (1997). *Writing test items to evaluate higher order thinking*. Allyn & Bacon.

Harden, R. M., Sowden, S., & Dunn, W. R. (1984). Educational strategies in curriculum development: The SPICES model. *Medical Education*, *18*(4), 284–297. doi:10.1111/j.1365-2923.1984.tb01024.x PMID:6738402

Kember, D. A. (1997). A reconceptualisation of the research into university academics' conceptions of teaching. *Learning and Instruction*, *7*(3), 255–275. doi:10.1016/S0959-4752(96)00028-X

Kremer, L., & Hofman, J. E. (1979). A Three-Dimensional Typology of Teacher Personality. *The Journal of Educational Research*, *73*(1), 20–25. doi:10.1080/00220671.1979.10885198

Kubiszyn, T., & Borich, G. (2013). *Educational testing and measurement*. John Wiley & Sons, Inc.

Kuntze, S. (2012). Pedagogical content beliefs: Global, content domain-related and situation-specific components. *Educational Studies in Mathematics*, *79*(2), 273–292. doi:10.100710649-011-9347-9

Mama, M., & Hennessy, S. (2013). Developing a typology of teacher beliefs and practices concerning classroom use of ICT. *Computers & Education*, *68*, 380–387. doi:10.1016/j.compedu.2013.05.022

Miller, M. D., Linn, R. L., & Gronlund, N. E. (2009). *Measurement and assessment in teaching*. Pearson.

Moskal, B. M. (2000). Scoring rubrics: What, when and how? *Practical Assessment, Research & Evaluation*, *7*(3), 1–5. doi:10.7275/a5vq-7q66

Nespor, J. (1987). The role of beliefs in the practice of teaching. *Journal of Curriculum Studies*, *19*(4), 317–328. doi:10.1080/0022027870190403

Nitko, A. J., & Brookhart, S. M. (2014). *Educational Assessment of Students* (6th ed.). Pearson.

Örücü, D. (2012). İlköğretim sınıf öğretmenlerinin sınıfa ve sınıf yönetimine ilişkin metaforik bakışları: Karşılaştırmalı bir durum çalışması. *İlköğretim Online, 11*(2), 342-352.

Popham, W. J. (2017). *Classroom assessment*. Pearson.

Rokeach, M. (1968). *Beliefs, attitudes and values: A theory of organization and change*. Jossey-Bass Inc.

Russell, M. K., & Airasian, P. W. (2012). *Classroom assessment*. McGraw-Hill.

Tenam-Zemach, M., & Flynn, J. E. (2015). *Rubric nation, critical inquiries on the impact of rubrics in education*. Information Age Publishing, Inc.

Törner, G. (2002). Mathematical beliefs-A search for a common ground: Some theoretical considerations on structuring beliefs, some research questions, and some phenomenological observations. In G. Leder, E. Pehkonen, & G. Törner (Eds.), *Beliefs: A Hidden Variable in Mathematics Education?* (pp. 73–94). Kluwer. doi:10.1007/0-306-47958-3_5

Vatansever, K. (2011). Öğretim üyelerinin öğretme yönelimlerinin belirlenmesi: Klinisyen öğretim üyeleriyle niteliksel bir çalışma (Identifying clinical teachers' orientation towards teaching: Qualitative survey of clinical teachers). *Tıp Eğitimi Dünyası*, *29*, 34–47.

Wubbels, T., & Levy, J. (1993). *Interpersonal relationships in education*. The Falmer Press.

Chapter 17

What Is Next for Rubrics?
A Reflection on Where We Are and Where to Go From Here

Heidi L. Andrade
University at Albany, USA

ABSTRACT

Rubrics have become ubiquitous in compulsory education and common in higher education. As with any educational innovation, it is time to reflect on the current state of rubrics and how to move ahead. This chapter identifies common conceptions of the rubric that are problematic and proposes redefining the word rubric in terms of learning goals to better align with classroom assessment uses. Feasible suggestions for ensuring the quality of rubrics and avoiding unintended negative consequences for students are also discussed.

INTRODUCTION

My first article about rubrics was published in *Educational Leadership* in 1996—nearly three decades ago. Since then, rubrics have become ubiquitous in compulsory education and common in higher education. As with any educational innovation, it is time to stand back and reflect on the current state of rubrics and how we might move ahead. In this chapter I focus on our conception of the rubric as problematic but salvageable and propose that we reframe the term to better align with classroom assessment uses.

BACKGROUND

It is tempting to begin with a brief history of rubrics in education, but Dawson (2017) has provided a detailed review that I will not attempt to improve upon except to add that the Rubric was part of the Catholic church's fourteenth-century Service Book, which described in excruciating detail the rules regarding religious ceremonies, vestments, the placement of furniture, and the like (Lewis, 1877). That is not what we are talking about here, obviously. The definition of the term *rubric* that I will defend in this

chapter is as follows: A rubric is a document that articulates the learning goals for a task and describes varying levels of mastery of those goals.

This is an unusual definition because of the reference to *learning goals* rather than *expectations* (e.g., Stevens & Levi, 2013) or *criteria* (e.g., Brookhart & Chen, 2015). It represents my attempt to reframe the definition of *rubric* to more explicitly represent its usefulness as a tool for teaching and learning. Rubrics have their roots in measurement, and much of the early research on them emphasized their summative aspects (Panadero & Jönsson, 2013). But the purposes of rubrics go far beyond measurement. The purposes of a rubric are to support teachers in designing instruction that addresses the goals, communicating the goals to students, guiding feedback from a variety of sources on students' progress toward the goals, and judging final products in terms of the degree to which the goals were met. Happily, in the hands of educators the rubric has come to serve the purposes of learning, but the measurement-oriented definition stuck. That is, we define rubrics in terms of grading and scoring, but we use them as instructional tools. We think of rubrics as useful tools for feedback, but we write rubrics as guidelines for evaluation. It is this disconnect between form and function, or features and purpose, that is the main concern of this chapter.

This disconnect is not surprising, given that at least some popular guidance on rubric design neglects learning goals (e.g., Panadero & Jönsson, 2020). In fact, the very definition of the term *rubric* often emphasizes measurement and grading, often to the exclusion of learning goals, objectives, or targets. I made that mistake myself back in 1996 when I defined a rubric as "a scoring tool that lists the criteria for a piece of work, or 'what counts'... [and] articulates gradations of quality for each criterion, from excellent to poor" (p. 14). I went on to add that "they are powerful tools for both teaching and assessment" and can "help students become more thoughtful judges of the quality of their own and others' work when they are used to guide self- and peer assessment" (p. 14), but those teaching and learning-oriented purposes were secondary to the scoring tool definition.

RUBRICS FOR MEASUREMENT VERSUS LEARNING

This measurement-based definition followed by learning-oriented uses is common. Mabry (1999) might call it a problem of "the shackles of psychometric habit" (p. 678). Take, for example, this definition from an excellent article by McTighe and Frontier (2022) about using rubrics to guide feedback: "Rubrics are typically used by teachers to judge the degree of students' understanding, proficiency levels of skills, the quality of their products or performances, and their growth from one level to the next. But beyond being evaluation tools, rubrics can be an excellent way to give feedback for improving teaching and learning" (p. 17). Judgment and measurement first, teaching and learning second. For another example, take Dawson's (2017) definition. After pointing out that the term has a variety of meanings, he writes, "a rubric is a tool used in the process of assessing student work that usually includes Popham's (1997) three essential features: evaluative criteria, quality definitions for those criteria at particular levels and a scoring strategy" (p. 349). This is an evaluative definition but like me (1996), McTighe and Frontier (2022), and many others (e.g., Arter & McTighe, 2001; Smit et al., 2017), he goes on to list their instructional uses as well: "Rubrics are more than just a tool used to support assessors in making summative judgements. Teachers also use rubrics as a way to provide feedback information.... Students can use rubrics in a range of ways, including self and peer assessment, and in interrogating the requirements of a task" (p. 355).

By now you might be thinking, so what? Why does this matter? I believe that it matters because the measurement-first conception of rubrics has led to a focus on the evaluative design and content of rubrics themselves, rather than how to write them to support learning—we *use* rubrics for instruction and feedback, but we *write* them like tests. As a result, students interpret them as tests or lists of requirements to which they must comply, which can lead to shallow learning and poor performances (Brookhart, 2013; Brookhart & Chen, 2015; Mabry, 1999; Torrance, 2007).

Aligning rubrics with learning goals or objectives is not a new idea. Biggs (1999) advocated a clear alignment between objectives, teaching and learning activities and assessment tasks. More recently, Chappuis et al. (2012) and Brookhart (2013) frame rubrics as guides for students to understand learning targets and monitor their own progress toward them. That is, rubrics can and should assess learning instead of task completion. And yet, many rubrics fail to do so often enough to elicit a strong warning from a leading voice in the field: "To co-opt rubrics into quality rating scales does violence, in my mind, to the whole point and purpose of using rubrics in the first place.... Be on the lookout for those and stamp out their use whenever possible. They are Trojan horses that will allow old-fashioned grading judgments to slip in where rubrics were intended" (Brookhart, 2013, p. 81). Popham also warned us back in 1997, when he warned that many rubrics were not instructionally beneficial.

I became aware of how little progress we have made regarding this issue while working with a group of arts educators, many of whom actively wrestled with the tension between rubrics as grading tools and rubrics as teaching tools. A conversation with a visual arts teacher and her professional learning group illustrates the tension. The art teacher reported that she was disappointed by the low, uninspired quality of her students' work on a collage project. She wisely suspected that her rubric was part of the problem and asked the group for suggestions. We homed in on a criterion for the use of paper in the collage, which referred to the requirement to use "at least three kinds of paper."

"What are your learning goals for this project?" we asked. "Do you really just want them to learn how to use three kinds of paper?"

"Of course not," she replied. "I want them to make paper choices that enhance the textural and expressive subject of the collage."

"That's what your rubric should say," we nearly shouted.

"But can you grade that?"

That question prompted a conversation about how arts educators critique art all the time and, when the goals of a project are clear, they can distinguish between student work that does and does not demonstrate mastery of those goals. We also talked about how useful rubrics can be in teaching the concepts and vocabulary of the discipline to students, as well as encouraging revision and creative risk-taking. The art teacher agreed to revise her rubric and try again. She later reported that, once the rubric focused on the intended learning, the students' work was elevated.

The visual arts, music, theater and dance teachers I work with often remark that their students come to talk and think like artists when their rubrics use the language of the arts and are used to guide peer feedback and self-assessment on their work (Andrade et al., 2019). We also have rigorous research evidence on improved performance-based learning outcomes in all four art forms (Chen & Andrade, 2016;

Chen et al., 2017; Mastrorilli et al., 2014). The rubrics used with, on, and by the students in that research were designed to serve formative functions by some of those same arts teachers who wrestled with the summative-formative tension in that conversation. I attribute their success to their focus on learning in the arts, which helped them make rubrics and checklists with criteria that aligned to their learning goals, not just the task. As a result, the teachers could do what Bearman and Ajjawi (2019) believe they should do, which is to convey teachers' intentions while also prompting students to develop their own ways of working and learning.

Breaking the shackles of psychometric habit is worth the effort. Rubrics designed to serve summative purposes are very different from those designed to serve formative purposes, as they should be. Take, for example, a quintessentially summative rubric formerly used by the College Board to score essays written for the recently discontinued Scholastic Aptitude Test (SAT) essay subject test. The SAT is an entrance exam used by many American colleges and universities to make admissions decisions. For the purposes of this illustration, I excerpted the sections of the rubric that were used to score cohesiveness and use of language:

4: Is cohesive and demonstrates a highly effective use and command of language.
3: Is mostly cohesive and demonstrates effective use and control of language.
2: Demonstrates little or no cohesion and limited skill in the use and control of language.
1: Demonstrates little or no cohesion and inadequate skill in the use and control of language.

That rubric was used to score hundreds of thousands of essays, but it clearly has little value in terms of teaching or learning about writing. I am not claiming that it should—it was designed to measure, not to instruct—but I have seen it reproduced in rubrics used in classroom contexts, where it should be refocused on the learning goals. Such a rubric might look something like this three-level description of cohesion excerpted from a rubric used for diagnostic assessment:

3: Within paragraphs, the individual sentences are seamlessly linked together; the reader can see the relationship between the ideas or information in one sentence and those in another sentence. The writing explicitly links sentences and ideas using adverbs (e.g., similarly, also, therefore), relative pronouns (e.g., who, that, which), conjunctions (e.g., and, or, while, whereas), and/or the repetition of key words, as appropriate.
2: The ideas or information in each sentence within a paragraph are generally but not consistently linked together, if only loosely. Additional or better choices of linking words and phrases would clarify the connections b/w ideas within paragraphs.
1: The connections between ideas in sentences within paragraphs are unclear. Little effective use of linking words and phrases.

Those descriptions of cohesion are from a rubric my research team and I wrote for a diagnostic writing assessment explicitly designed to provide students with feedback and resources for improving their writing (https://my.daacs.net/). I hope it is obvious that I am not saying one rubric is better than the other, since they are not really comparable: The former is for measurement, and the latter is for learning. My point is that the features of our rubrics should more explicitly serve their functions: The SAT rubric served the purposes of measurement and evaluation, while the DAACS rubric serves the purposes of

feedback. Rubrics with learning at the front-end invite students to build and demonstrate mastery of the goals of the task.

HOW DO WE GET THERE? RECOMMENDATIONS FOR PRACTICE

Refocusing rubrics used in classroom contexts on learning has implications for practice. In this section of the chapter, I make several proposals that, given the current state of teacher literacy regarding classroom assessment (Pastore & Andrade, 2019), will have implications for both pre- and in-service teacher education. Given all my talk in this chapter about throwing off the shackles of psychometric habit, it might be surprising to read that there is one such habit that I think would assist us in writing better rubrics: the habit of reflecting on evidence of validity, reliability, and fairness.

Validity and Reliability

Rubrics used in low-stakes classroom contexts are not exempt from the demands of validity and reliability (Moskal & Leydens 2000; Tierney, 2022). Validity and reliability are concerned with the consistency and accuracy of the judgments we make about students and their work (Payne, 2003). At a minimum, a rubric must be aligned with reasonable standards that are grounded in the discipline and with the curriculum being taught in order to support valid conclusions about student learning. It must pass a test of reliability by resulting in similar ratings when used by different people. These concerns do not require us to perform complex statistical analyses but, rather, that we worry enough about them to design our rubrics carefully and subject them to critique.

Rubrics improve when we compare them to standards and learning goals, show them to another teacher, or ask a colleague to co-score some student work. Payne (2003) suggests that sitting and listening to students critique our assessments and the associated rubrics can be the best source of information about how good they are, and I agree (Goodrich Andrade, 2005). Let us consider a few essential steps that all teachers in primary, secondary, and higher education can reasonably be expected to take when creating or adapting and using a rubric.

Check Alignment With Learning Goals

Articulating goals or objectives and describing them in rubrics is hard work—a nontrivial task that requires considerable skill (Sadler, 2005). According to Bacchus et al. (2019), part of the problem is that criteria and learning outcomes are often developed during separate processes. I strongly recommend that anyone writing a new rubric or adapting an existing one begins by explicitly checking the alignment of criteria and learning goals. In this way, learning goals are drawn into the process of creating a rubric, and the likelihood of alignment is increased. A simple table like Table 1 can support alignment checking.

Table 1. Checking alignment between learning goals and rubric criteria

Task-relevant Learning Goals	Related Criterion/a

Once the criteria and goals are aligned, a rubric should be evaluated in terms of the criteria in Table 2, which was adapted from Brookhart (2013). Table 2 lists the criteria for high quality rubrics and a brief description of each; see Brookhart's book for more detail. I use Tables 1 and 2 in my graduate course on classroom assessment to scaffold student self-assessment of their original rubrics. Having used Table 1 to check for alignment with the learning goals, they can check off the first and arguably most important criterion in Table 2: Appropriate.

Table 2. Criteria for rubrics

The <u>criteria</u> are…	
Appropriate	Each criterion is related to a learning goal. The criteria answer the question, "What characteristics of student work would provide evidence of student learning?"
Observable	Each criterion describes a quality of student work that can be seen or heard. (This disqualifies "effort" as a criterion.)
Distinct	Each criterion identifies a separate aspect of the learning outcomes the task is intended to teach and assess.
Complete	Taken together, the criteria comprehensively describe the learning goals the task is intended to assess.
The <u>levels</u> are…	
Descriptive	Student work is described, not counted.
Clear	Students and teachers can understand what the descriptions mean.
Cover the whole range	Student work is described from one extreme to the other for each criterion
Distinctive	The descriptions are different enough from level to level to categorize work unambiguously (within reason).
Parallel	The descriptions at each level describe different qualities for the same aspects of student work.

Adapted from Brookhart, S. M. (2013). *How to create and use rubrics for formative assessment and grading*. Association for Supervision & Curriculum Development.

Check for Clarity and the Quality of Language

Another of Brookhart's (2013) criteria for good rubrics is *Clear*. Unfortunately, many rubrics are what Leisen (2022) refers to as a patchwork quilt of teacher-talk, or a wall of words. A likely result of a wall of words is confusion on the part of students who, research by Bacchus et al. (2019) has shown, stop reading them:

While 71% of respondents said they always read the rubrics, only 43% said they usually understood what was required from reading rubrics. The purpose of rubrics is to be used as a learning tool for completing assessments, yet these results suggest that they are meeting this aim in fewer than 50% of cases. In conjunction with the finding that some of those who did not read the rubrics did not do so because they believed the rubrics were too complicated and confusing, this strongly suggests many students do not benefit from the provision of rubrics per se. (p. 4)

Bacchus et al. conclude that students need instruction and experience to learn to interpret and make the most of rubrics. I agree but also want to stress the importance of providing clear rubrics that use accessible, growth-oriented language. In addition to the quantity of words on a rubric, it is important to get the quality of those words right. I learned a lot about this from my colleague, Peter Johnston (2004; Johnston & Andrade, 2012), who wrote extensively about the powerful influence of teachers' and parents' word choices on children's intellectual, relational, epistemological, and even moral agency. Peter taught me that the language of rubrics should be informed by theory and research on mindset (Dweck, 1999).

Take, for example, a study by Cimpian and his colleagues, including Dweck (2007) in which they had young children use puppets to role-play drawing three pictures and then had a puppet-teacher give them feedback. The puppet-teacher deemed the first three drawings successful, and gave half the children the response, "You're a good drawer." The other half received the feedback, "You did a good job drawing." The difference is subtle: "You're a good drawer" frames the activity as one in which the child is being judged as an artist and a person. The "good job drawing" feedback frames the activity as one in which the quality of the process of doing the work is judged, but not the person.
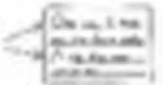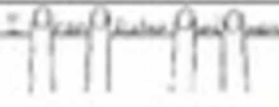
After three instances of one or other of these forms of positive feedback, the children were asked a series of questions about the extent to which they liked a drawing they had done, whether they felt happy or sad, good or not good, and whether they felt like a good or not good person. There was no difference between the groups. During the next two role-play events, both groups received non-judgmental feedback suggesting that their drawing had been unsuccessful, such as, "The cat [you just drew] has no ears." Asked the same set of questions again, now about their unsuccessful drawings, the children who had been induced to think that the activity was about deciding who was and was not a good artist became more negative in their responses. Asked which picture they might draw on another day, these students chose one that they had already successfully drawn. Asked whether they would choose to draw or do something else, they were more likely to choose to do something else. The children who had been led to believe the activity was about the process of doing a good job responded more positively in every way and were more likely to choose the more challenging activity on which they had so far been unsuccessful.

The study shows how the language used in assessment mediates the meaning of the children's experience of the activity. The two groups of children ostensibly had the same experience but were induced to view it and its significance differently. The different framing affected the students' goals, teacher and student roles, the relationship between teacher and student, and the relationship between the student and the activity. Language that frames the outcome of an assessment as evidence of a permanent trait routinely leads to subsequent avoidance of challenge (Dweck, 1999). The qualities of the words we include in our rubrics potentially have an impact, for better or worse, on subsequent learning, and not only on the content being learned, but also on broader aspects of children's development.

In light of this perspective, consider the frequent use of smiling and frowning faces in rubrics like this one, which I intentionally do not cite (Figure).

Figure 1. K-2 "writing checklist"

This rubric is used with five- and six-year-old budding writers. It judges rather than instructs. It shames rather than guides. I imagine its author chose to use emojis to accommodate preliterate students, but the intense frowns can have unintended consequences for children's learning and their relationships with their teachers that are easily avoided if we write with growth-oriented language and images in mind.

In addition to avoiding frowning or even crying at students (I have seen it), authors of rubrics should also be clear that they are assessing student work, not students themselves. The difference between "the student cannot use math vocabulary" and "the explanation does not use math vocabulary" are probably self-evident.

For a contrasting example, see Maria Comba's fourth-grade melody rubric in Table 3 (Valle et al., 2016). Ms. Comba is an elementary school music teacher in Brooklyn, New York. She designed a unit to teach students to hear melody lines and notate simple melodies. The unit focused on ear training and melodic dictation. The learning goals for the lesson were to: 1) understand the concept of melody, 2) distinguish between the melody line and the accompaniment, 3) understand how melody is developed, 4) use vocabulary appropriately when speaking about melody. At the beginning of the unit, students were asked to listen carefully to a melody in order to draw the shape, or contour, of the melody line and use this rubric to self-assess.

Table 3. Maria Comba's fourth grade melody rubric

	I've GOT it!	**Ah-ha … I'm almost there!**	**I'm getting better!**	**I need some help, please.**
Drawing the Contour	I've got it! Not only can I draw the contour, but I can add details so it starts to look like a melody on the staff.	Ah-ha! I can draw the contour correctly. ***Now I'm working on*** notating some details so it can start to look like a melody line on a staff.	I can draw the contour when it moves in one direction. ***Now I'm working on*** "listening for movement" in contours that move in different directions.	***Now I'm working on*** "listening for movement." I'm always asking myself if the music sounds like it's moving up the stairs or down the stairs.
Notating the Melody on the Staff	I can notate the melody when given the starting pitch. I used "listening for movement," melodic motion (ascending, descending, step, skip), and solfège to help me.	I can use the starting pitch and melodic motion to figure out the movements and relationships of pitches in the melody. ***Now I'm working on*** writing notes closer to their actual pitch. Using melodic motion and solfège will help.	I can notate the direction of the contour by using melodic motion but cannot place them on the staff as of yet. ***Now I'm working on*** using the starting pitch to help me place notes on the staff.	***Now I am working on*** "listening for movement" and matching it up with melodic motion. I will always follow picture cues with my finger to see if they match.
Singing the Melody	I can sing the melody line when given the starting pitch. I used melodic motion and solfège to help me.	I can sing the melody line moving in the right direction but the pitches are not accurate. ***Now I'm working on*** being more accurate with each individual pitch.	I can sing the contour but did not match individual notes. ***Now I'm working on*** using solfège to help me sing the correct pitch.	***Now I'm working on*** making sure that my voice is going in the right direction. I'm listening, tracing contours and echoing.

The *now I'm working on* is your goal for the next class.

Notice how the repeated phrase "now I'm working on" helps students identify specific areas in need of improvement, goals for their work, and strategies for meeting those goals. It also avoids negative phrasing of rubric descriptors. The explicitly stated goals and next steps on the rubric give students immediate and appropriate tactics for deepening their learning. Ms. Comba took a growth-oriented, learning-first approach to writing this rubric. There is reason to believe that makes a difference in terms of student learning as well as the development of self-regulated learning skills and dispositions such as goal-setting and self-assessment (Andrade & Heritage, 2017).

Check for Reliability

A common way to check for consistency in grading with a rubric is to go back over a graded set of papers to ensure that assessment criteria is evenly applied, which means the scores were not affected by mood, fatigue, or implicit bias (Tierney, 2013). The suggestion to do so is usually met with shock by teachers, who have a hard enough time grading all student work *once,* much less twice. Instead, I recommend re-grading just a few pieces of student work, rather than all of it. I do this myself, and it has helped me make improvements to my rubrics. I also, on occasion, ask another expert in my field to grade a few of the most important (that is, mostly heavily weighted) pieces of student work. When we disagree on a score for a particular criterion, we talk it over and determine whether the rubric can be revised to avoid disagreements. My rubrics are invariably clearer after this process – to me and to my students.

Teacher teams and professional learning communities are appropriate contexts for enhancing the validity and reliability of rubrics used for classroom assessment. In addition, professional development for

teachers should stress standards and learning goals as the starting point for writing rubrics. Coursework and texts used in teacher preparation classes should include rubric development and use as pedagogy, not just measurement. One of my graduate students reinforced these points for me recently when she relayed a conversation with her supervisor, the Director of Academic Services at a local community college. The Director thought it was silly that my student was studying rubric development and said, "When I was an adjunct, I would slap some criteria on a rubric and do quick math. It isn't that difficult or complex". Whether or not rubric design is difficult or complex, it is important to get it right, given the consequences for student achievement. I urge instructors at all levels of the educational system to take the steps described above to enhance the quality of these ubiquitous assessment tools.

DIRECTIONS FOR RESEARCH

There are many interesting studies to be done. Here I will mention just three. First, I have stressed the importance of learning-first rubrics in this chapter but whether or not they make a difference is an open empirical question. I would love to see a carefully designed study that compares a measurement-oriented rubric with a learning-goal-oriented rubric. The study should employ a fine-grained measure of learning.

Also, given claims about the potential for rubric use to promote student self-regulated learning (SRL; Andrade & Heritage, 2017; Panadero et al., 2019), it would also be valuable to address questions about whether learning-goal-oriented rubrics are associated with learning-oriented SRL strategy use. A task-specific measure of SRL with evidence of validity should be used for such a study.

Finally, since students tend to think of rubrics as learning tools while teachers often perceive of them as tools for quickly, objectively, and accurately assigning grades (Li & Lindsey, 2015; Parkes, 2006), research that brings the two views in alignment is welcomed. I imagine a qualitative study in which students teach teachers about how and why they use rubrics, and the features of the rubrics that serve them best. The effects of such discussions on the rubrics teachers produce and the ways in which they employ them could be worth capturing.

CONCLUSION

Teachers, schools and districts have embraced rubrics and consider them as useful tools for feedback as well as grading (e.g., see Garcez-Manzanera, 2022). Research shows that, under the right conditions, they can promote learning, effective peer and self-assessment, and even self-regulated learning, while also expediting the marking process and indicating whether and where instruction fell short (Andrade & Heritage, 2017; Brookhart & Chen, 2015; Camargo et al., 2023; Jönsson & Svingby, 2007; Panadero & Jönsson, 2013). Given their promise, I challenge us and others to

- think of rubrics not primarily as instruments for judging but as tools for teaching;
- rather than defining rubrics as scoring guides, define them as information about the learning goals for a task with descriptions of varying levels of mastery of those goals;
- highlight the teaching and learning purposes of rubrics, which include supporting teachers in designing instruction that addresses the goals, communicating the goals to students, guiding feed-

back from a variety of sources on students' progress toward the goals, and judging final products in terms of the degree to which the goals were met.

- be careful to avoid images (frown faces) and language ("the student cannot") that we know are likely to undermine students' achievement, attitudes toward learning, and relationships with their teachers; and

- solicit feedback on original or adapted rubrics from other teachers and students and be prepared to revise them to ensure they articulate the learning goals for the task and can be used to reliably grade student work.

REFERENCES

Andrade, H., Hefferen, J., & Palma, M. (2019). Formative assessment in the arts. In H. Andrade, R. Bennett, & G. Cizek (Eds.), *Handbook of formative assessment in the disciplines* (pp. 126–145). Routledge. doi:10.4324/9781315166933-6

Andrade, H., & Heritage, M. (2017). *Using assessment to enhance learning, achievement, and academic self-regulation*. Routledge. doi:10.4324/9781315623856

Arter, J., & McTighe, J. (2001). *Scoring rubrics in the classroom: Using performance criteria for assessing and improving student performance*. Corwin Press.

Bacchus, R., Colvin, E., Bronwen Knight, E., & Ritter, L. (2019). When rubrics aren't enough: Exploring exemplars and student rubric co-construction. *Journal of Curriculum and Pedagogy*. doi:10.1080/15505170.2019.1627617

Bearman, M., & Ajjawi, R. (2019). Can a rubric do more than be transparent? Invitation as a new metaphor for assessment criteria. *Studies in Higher Education*, *46*(2), 359–368. doi:10.1080/03075079.2019.1637842

Biggs, J. (1999). *Teaching for quality learning at university: what the student does*. Open University Press.

Brookhart, S., & Chen, F. (2015). The quality and effectiveness of descriptive rubrics. *Educational Review*, *67*(3), 343–368. doi:10.1080/00131911.2014.929565

Brookhart, S. M. (2013). *How to create and use rubrics for formative assessment and grading*. Association for Supervision & Curriculum Development.

Camargo, S., Chang, A., Maeda, Y. (2023, April). *Effect of rubrics on motivation and performance: Meta-analysis study.* [Paper Presentation] American Educational Research Association.

Chappuis, J., Stiggins, R. J., Chappuis, S., & Arter, J. (2012). *Classroom assessment for student learning: Doing it right-using it well*. Pearson.

Chen, F., & Andrade, H. (2016). The impact of criteria-referenced formative assessment on fifth grade students' theater arts achievement. *The Journal of Educational Research*, *109*, 1–10.

Chen, F., Lui, A., Andrade, H., Valle, C., & Mir, H. (2017). Criteria-referenced formative assessment in the arts. *Educational Assessment, Evaluation and Accountability*, *29*(3), 297–314. doi:10.100711092-017-9259-z

Cimpian, A., Arce, H.-M. C., Markman, E. M., & Dweck, C. S. (2007). Subtle linguistic cues affect children's motivation. *Psychological Science*, *18*(4), 314–316. doi:10.1111/j.1467-9280.2007.01896.x PMID:17470255

Dawson, P. (2017). Assessment rubrics: Towards clearer and more replicable design, research and practice. *Assessment & Evaluation in Higher Education*, *42*(3), 347–360. https://doi-org.libproxy.albany.edu/10.1080/02602938.2015.1111294. doi:10.1080/02602938.2015.1111294

Dweck, C. S. (1999). *Self-theories: their role in motivation, personality, and development*. Psychology Press.

Fraile, J., Panadero, E., & Pardo, R. (2017). Co-creating rubrics: The effects on self-regulated learning, self-efficacy and performance of establishing assessment criteria with students. *Studies in Educational Evaluation*, *53*, 69–76. doi:10.1016/j.stueduc.2017.03.003

Garcez-Manzanera, A. (2022). The affordances of rubrics in L2 writing in higher education: A new approach to enhancing writing conventions. *HUMAN Review*, *11*(Monográfico), 2–12. doi:10.37467/revhuman.v11.4086

Gikandi, J. W., & Morrow, D. (2016). Designing and implementing peer formative feedback within online learning environments. *Technology, Pedagogy and Education*, *25*(2), 153–170. doi:10.1080/1475939X.2015.1058853

Goodrich, H. (1996). Understanding rubrics. *Educational Leadership*, *54*(4), 14–17.

Johnston, P., & Andrade, H. (2012). Assessment, teaching and learning in and beyond classrooms. In B. Kaur (Ed.), *Understanding teaching and learning: Classroom research revisited* (pp. 269–280). Sense Publishers. doi:10.1007/978-94-6091-864-3_20

Johnston, P. H. (2004). *Choice words: How our language affects children's learning*. Stenhouse Publishers.

Jönsson, A., & Svingby, G. (2007). The use of scoring rubrics: Reliability, validity and educational consequences. *Educational Research Review*, *2*(2), 130–144. doi:10.1016/j.edurev.2007.05.002

Leisen, M. (2022). Make your rubric more than a wall of words. *Educational Leadership*, *79*(7). https://www.ascd.org/el/articles/make-your-rubric-more-than-a-wall-of-words

Lewis, J. (1877). The rubric: its history and meaning. In *The Oxford Movement: Tractarian Pamphlets at Pusey House: The Halifax and Church Sub-Collections* (2nd ed.). Adams, and Co.

Li, J., & Lindsey, P. (2015). Understanding variations between student and teacher application of rubrics. *Assessing Writing*, *26*, 67–79. doi:10.1016/j.asw.2015.07.003

Mabry, L. (1999). Writing to the rubric: Lingering effects of traditional standardized testing on direct writing assessment. *Phi Delta Kappan*, *80*(9), 673–679.

Mastrorilli, T. M., Harnett, S., & Zhu, J. (2014). *Arts Achieve* impacting student success in the arts: Preliminary findings after one year of implementation. *Journal for Learning through the Arts, 10*(1). https://escholarship.org/uc/item/6c81239d

McTighe, J., & Frontier, T. (2022). How to provide better feedback through rubrics. *Educational Leadership*, *79*(7), 17–23.

Panadero, E., Broadbent, J., Boud, D., & Lodge, J. (2019). Using formative assessment to influence self- and co-regulated learning: The role of evaluative judgement. *European Journal of Psychology of Education*, *34*(3), 535–557. doi:10.100710212-018-0407-8

Panadero, E., & Jönsson, A. (2013). The use of scoring rubrics for formative assessment purposes revisited: A review. *Educational Research Review*, *9*, 129–144. doi:10.1016/j.edurev.2013.01.002

Panadero, E., & Jönsson, A. (2020). A critical review of the arguments against the use of rubrics. *Educational Research Review*, *30*, 100329. doi:10.1016/j.edurev.2020.100329

Parkes, K. A. (2006). *The effect of performance rubrics on college level applied studio grading*. [PhD dissertation. University of Maimi]. UMI No. 3215237.

Pastore, S., & Andrade, H. L. (2019). Teacher assessment literacy: A three-dimensional model. *Teaching and Teacher Education*, *84*, 128–138. doi:10.1016/j.tate.2019.05.003

Popham, W. J. (1997). What's wrong – and what's right – with rubrics. *Educational Leadership,* Sadler, R. D. (2005). Interpretations of criteria-based assessment and grading in higher education. *Assessment & Evaluation in Higher Education*, *30*(2), 175–194. doi:10.1080/0260293042000264262

Smit, R., Bachmann, P., Blum, V., Birri, T., & Hess, K. (2017). Effects of a rubric for mathematical reasoning on teaching and learning in primary school. *Instructional Science*, *45*(5), 603–622. doi:10.100711251-017-9416-2

Stevens, D., & Levi, A. (2013). *Introduction to rubrics: An assessment tool to save grading time, convey effective feedback, and promote student learning* (2nd ed.). Stylus Publishing.

Tierney, R. D. (2013). Fairness in classroom assessment. In J. H. McMillan (Ed.), *SAGE handbook of research on classroom assessment* (pp. 125–144). SAGE Publications. doi:10.4135/9781452218649.n8

Tierney, R. D. (2022). Fairness in educational testing and assessment. In D. Fisher (Ed.), *Routledge encyclopedia of education*. Routledge. doi:10.4324/9781138609877-REE35-1

Torrance, H. (2007). Assessment as learning? How the use of explicit learning objectives, assessment criteria and feedback in post-secondary education and training can come to dominate learning. *Assessment in Education: Principles, Policy & Practice*, *14*(3), 281–294. doi:10.1080/09695940701591867

Valle, C., Andrade, H., Palma, M., & Hefferen, J. (2016). Applications of peer and self-assessment in music education. *Music Educators Journal*, *102*(4), 41–49. doi:10.1177/0027432116644652

Vasileiadou, D., & Karadimitriou, K. (2021). Examining the impact of self-assessment with the use of rubrics on primary school students' performance. *International Journal of Educational Research Open*, *2*, 100031. doi:10.1016/j.ijedro.2021.100031

Compilation of References

AACSB. (2019). Accreditation Standard 8 (2013 Business Standards): Curricula Management and Assurance of Learning An Interpretation, AACSB. https://www.aacsb.edu/insights/briefings/standard-8-white-paper

AACSB. (2020). 2020 Guiding Principles and Standards For Business Accreditation. https://www.aacsb.edu/educators/accreditation/business-accreditation/aacsb-business-accreditation-standards

Aberg-Bengtsson, T., & Ottosson, T. (2006). What lies behind graphicacy? Relating students' results on a test of graphically represented quantitative information to formal academic achievement. *Journal of Research in Science Teaching*, *43*(1), 43–62. doi:10.1002/tea.20087

Abrami, P. C., & Barrett, H. (2005). Directions for research and development on electronic portfolios. *Canadian Journal of Learning and Technology*, *31*(3), 1–15. doi:10.21432/T2RK5K

Accurso, K., & Gebhard, M. (2021). SFL praxis in US teacher education: A critical literature review. *Language and Education*, *35*(5), 402–428. doi:10.1080/09500782.2020.1781880

Adams, T. E., Ellis, C., & Jones, S. H. (2017) Autoethnography. The international encyclopaedia of communication research methods, pp.1-11.

Adarkwah, M. A. (2021). The power of assessment feedback in teaching and learning: A narrative review and synthesis of the literature. *SN Social Sciences*, *1*(3), 75. doi:10.100743545-021-00086-w

Adcroft, A. (2011). The mythology of feedback. *Higher Education Research & Development*, *30*(4), 405–419. doi:10.1080/07294360.2010.526096

Advance H. E. (2016) *Student Engagement Through Partnership Framework*. Advance HE.

Advance H. E. (2023), *The UK Engagement Survey (UKES)*. Advance HE. https://www.advance-he.ac.uk/reports-publications-and-resources/student-surveys/uk-engagement-survey-ukes

Afdal, H. W., & Spernes, K. (2018). Designing and redesigning research-based teacher education. *Teaching and Teacher Education*, *74*, 215–228. doi:10.1016/j.tate.2018.05.011

Agu, N. N., Omenyi, A. S., & O, C. (2015). Evaluation of doctorate dissertation in Nigerian Universities: do faculties provide and use criteria / rubrics? *International Journal of Technology and Inclusive Education (IJTIE), 4*(1), 565-569.

Aguirre-Muñoz, Z., Park, J.-E., Amabisca, A., & Boscardin, C. K. (2009). Developing teacher capacity for serving ELLs' writing instructional needs: A case for systemic functional linguistics. *Bilingual Research Journal*, *31*(1-2), 295–322. doi:10.1080/15235880802640755

Akar, H., & Yildirim, A. (2009). Change in teacher candidates' metaphorical images about classroom management in a social constructivist learning environment. *Teaching in Higher Education*, *14*(4), 401–415. doi:10.1080/13562510903050152

Aktaş, M., & Alıcı, D. (2018). Analytical rubric development for story writing: Validity and reliability study. *Mersin University Journal of the Faculty of Education*, *14*(2), 597–610. doi:10.17860/mersinefd.424198

Allen, D., & Tanner, K. (2006). Rubrics: Tools for Making Learning Goals and Evaluation Criteria Explicit for Both Teachers and Learners. *CBE Life Sciences Education*, *5*(3), 197–203. doi:10.1187/cbe.06-06-0168 PMID:17012210

Allen, S., & Knight, J. (2009). A Method for Collaboratively Developing and Validating a Rubric. *International Journal for the Scholarship of Teaching and Learning*, *3*(2), 3. doi:10.20429/ijsotl.2009.030210

Almeida, F., & Buzady, Z. (2019). Assessment of entrepreneurship competencies through the use of FLIGBY. *Digital Education Review*, 151-169. doi:10.1344/der.2019.35.151-169

Alsina, Á., Ayllón, S., & Colomer, J. (2019). Validating the Narrative Reflection Assessment Rubric (NARRA) for reflective narratives in higher education. *Assessment & Evaluation in Higher Education*, *44*(1), 155–168. doi:10.1080/02602938.2018.1486391

Alsina, Á., Ayllón, S., Colomer, J., Fernández-Peña, R., Fullana, J., Pallisera, M., Pérez-Burriel, M., & Serra, L. (2017). Improving and evaluating reflective narratives: A rubric for higher education students. *Teaching and Teacher Education*, *63*, 148–158. doi:10.1016/j.tate.2016.12.015

Álvarez, J. (2005). *Evaluar para conocer, examinar para excluir*. Morata.

Andrade, H., & Du, Y. (2005). Student Perspectives on Rubric-Referenced Assessment. *Educational & Counseling Psychology Faculty Scholarship*. https://scholarsarchive.library.albany.edu/edpsych_fac_scholar/2

Andrade, H., & Du, Y. (2007). *Student Responses to Criteria Referenced Self-assessment. Educational Administration & Policy Studies Faculty Scholarship*. Scholar Archive. scholarsarchive.library.albany.edu/eaps_fac_scholar/1

Andrade, H. (2000). Using Rubrics to Promote Thinking and Learning. *Educational Leadership*, *57*(5), 13–18.

Andrade, H. G. (1997). Understanding rubrics. *Educational Leadership*, *54*(4), 14–17.

Andrade, H. G. (2000). Using rubrics to promote thinking and learning. *Educational Leadership*, *57*(5), 13–19.

Andrade, H. G. (2001). The effects of instructional rubrics on learning to write. *Current Issues in Education (Tempe, Ariz.)*, *4*(4). http://cie.ed.asu.edu/volume4/number4/

Andrade, H. G. (2005). Teaching with rubrics: The good, the bad and the ugly. *College Teaching*, *53*(1), 27–31. doi:10.3200/CTCH.53.1.27-31

Andrade, H. G., & Boulay, B. A. (2003). Role of Rubric-Referenced Self-Assessment in Learning to Write. *The Journal of Educational Research*, *97*(1), 21–30. doi:10.1080/00220670309596625

Andrade, H. G., & Du, Y. (2005). Student perspectives on rubric-referenced assessment. *Practical Assessment, Research & Evaluation*, *10*(3), 1–11.

Andrade, H. L., & Brookhart, S. M. (2020). Classroom assessment as the co-regulation of learning. *Assessment in Education: Principles, Policy & Practice*, *27*(4), 350–372. doi:10.1080/0969594X.2019.1571992

Andrade, H. L., Wang, X., Du, Y., & Akawi, R. L. (2009). Rubric-referenced self-assessment and self-efficacy for writing. *The Journal of Educational Research*, *102*(4), 287–302. doi:10.3200/JOER.102.4.287-302

Andrade, H., & Du, Y. (2005). Student perspectives on rubric-referenced assessment. *Practical Assessment, Research & Evaluation*, *10*(1), 3. doi:10.7275/g367-ye94

Andrade, H., Hefferen, J., & Palma, M. (2019). Formative assessment in the arts. In H. Andrade, R. Bennett, & G. Cizek (Eds.), *Handbook of formative assessment in the disciplines* (pp. 126–145). Routledge. doi:10.4324/9781315166933-6

Andrade, H., & Heritage, M. (2017). *Using assessment to enhance learning, achievement, and academic self-regulation.* Routledge. doi:10.4324/9781315623856

Andrade, H., & Valtcheva, A. (2009). Promoting Learning and Achievement Through Self-Assessment. *Theory into Practice*, *48*(1), 12–19. doi:10.1080/00405840802577544

Andrews, M., Brown, R., & Mesher, L. (2018). Engaging students with assessment and feedback: Improving assessment for learning with students as partners. *Practitioner Research in Higher Education Journal*, *11*(1), 32–46.

Anglin, L., Anglin, K., Schumann, P. L., & Kaliski, J. A. (2008). Improving the efficiency and effectiveness of grading through the use of computer-assisted grading rubrics. *Decision Sciences Journal of Innovative Education*, *6*(1), 51–73. doi:10.1111/j.1540-4609.2007.00153.x

Angra, A., & Gardner, S. M. (2018). The graph rubric: Development of a teaching, learning, and research tool. *CBE Life Sciences Education*, *17*(4), 1–18. doi:10.1187/cbe.18-01-0007 PMID:30496033

Aronson, L., Niehaus, B., Hill-Sakurai, L., Lai, C., & O'Sullivan, P. S. (2012). A comparison of two methods of teaching reflective ability in Year 3 medical students. *Medical Education*, *46*(8), 807–814. doi:10.1111/j.1365-2923.2012.04299.x PMID:22803758

Arpaguş, E. K., Ünsal, Y., & Moğol, S. (2011). The Effects of visual literacy on the success of secondary school students in spherical mirrors and lenses. [NWSA]. *E-Journal of New World Sciences Academy*, *6*(3), 1972–1981.

Arter, J. (2000). Rubrics, scoring guides, and performance criteria. Classroom tools for Assessing and improving student learning. The American Educational Research Association, New Orleans, LA, April, 1-22.

Arter, J. A., & Chappuis, J. (2007). *Creating & recognizing quality rubrics*. Pearson Merrill Prentice Hall.

Arter, J. A., & McTighe, J. (2001). *Scoring rubrics in the classroom: Using performance criteria for assessing and improving student performance*. Corwin Press.

Arter, J., & McTighe, J. (2001). *Scoring Rubrics in the Classroom: Using Performance Criteria for Assessing and Improving Student Performance*. SAGE Publications.

Ashton, S., & Davies, R. S. (2015). Using scaffolded rubrics to improve peer assessment in a MOOC writing course. *Distance Education*, *36*(3), 312–334. doi:10.1080/01587919.2015.1081733

Aslanoğlu, A. E., & Kutlu, Ö. (2003). Öğretimde sunu becerilerinin değerlendirilmesinde dereceli puanlama anahtarı (rubrıc) kullanılmasına ilişkin bir araştırma. *Ankara University*. *Journal of Educational Sciences*, *36*(1), 25–36.

Athayde, R. (2009). Measuring enterprise potential in young people. *Entrepreneurship Theory and Practice*, *33*(2), 481–500. doi:10.1111/j.1540-6520.2009.00300.x

Athon, A. (2019). Designing Rubrics to Foster Students' Diverse Language Backgrounds. *Journal of Basic Writing*, *38*(1), 78–104. doi:10.37514/JBW-J.2019.38.1.05

Atkinson, D., & Lim, S. L. (2013). Improving assessment processes in Higher Education: Student and teacher perceptions of the effectiveness of a rubric embedded in a LMS. *Australasian Journal of Educational Technology*, *29*(5), 651–666. doi:10.14742/ajet.526

Austin, A. E. (2002). Preparing the next generation of faculty: Graduate school as socialization to the academic career. *The Journal of Higher Education*, *73*(1), 94–122. doi:10.1080/00221546.2002.11777132

Auxtero, L. C., & Callaman, R. A. (2020). Rubric as a learning tool in teaching application of derivatives in basic calculus. [Journal of Research and Advances in Mathematics Education]. *JRAMathEdu*, *6*(1), 46–58. doi:10.23917/jramathedu.v6i1.11449

Avraamidou, L. & ve Zembal-Saul, C. (2006). Exploring the influence of web-based portfolio development on learning to teach elementary science. *AACE Journal*, *14*(2), 178–205.

Aydın Gürler, S., & Baykara, O. (2020). The use of scoring rubric in science course and student opinions on the scoring rubric: A mixed method study. *Turkish Studies -. Education in Science*, *15*(2), 673–690.

Aydın, A., & Tarakçı, F. (2018). The Investigation of the pre-service science teachers' abilities to read, interpret and draw graphs. *Elementary Education Online, 17*(1), 469-488. https://doi.org/. 2018.413806 doi:10.17051/ilkonline

Aydın, N. (2018). Determining of science pre-service teachers' graphic understanding and interpretation levels in context of heat and temperature. *Fen Bilimleri Öğretimi Dergisi*, *6*(1), 20–36.

Ayhan, Ü., & Türkyılmaz, M. U. (2015). Key of language assessment: Rubrics and rubric design. *International Journal of Language and Linguistics*, *2*(2), 82–92.

Ayre, C., & Scally, A. J. (2014). Critical values for Lawshe's content validity ratio: Revisiting the original methods of calculation. *Measurement & Evaluation in Counseling & Development*, *47*(1), 79–86. doi:10.1177/0748175613513808

Babatunde, S., & El-Gohary, H. (2019). Necessity of mentoring in entrepreneurship education: Reflection by practitioners. *Journal of Professional Issues in Engineering Education and Practice*, *145*(1), 02518007. doi:10.1061/(ASCE)EI.1943-5541.0000399

Babatunde, S., El-Gohary, H., & Edwards, D. (2021). Assessment methods in entrepreneurship education, challenges and opportunities in developed and developing nations: A comparative study of Nigeria and England. *Education + Training*, *63*(7/8), 1092–1113. doi:10.1108/ET-12-2020-0368

Bacchus, R., Colvin, E., Bronwen Knight, E., & Ritter, L. (2019). When rubrics aren't enough: Exploring exemplars and student rubric co-construction. *Journal of Curriculum and Pedagogy*. doi:10.1080/15505170.2019.1627617

Bache, C. (2013). Grammatical choice and communicative motivation: a radical systemic approach. In *Systemic functional linguistics: Exploring choice* (pp. 72–94). Cambridge University Press. doi:10.1017/CBO9781139583077.006

Bacigalupo, M., Kampylis, P., Punie, Y., & Van den Brande, G. (2016). EntreComp: The entrepreneurship competence framework. *Luxembourg. Publication Office of the European Union*, *10*, 593884.

Bahtaji, M. A. A. (2020). Improving students graphing skills and conceptual understanding using explicit graphical physics instructions. *Cypriot Journal of Educational Science.*, *15*(4), 843–853. doi:10.18844/cjes.v15i4.5063

Bailey, R.A., (2009). Undergraduate students' perceptions the role and utility of written assessment feedback. *Journal of Learning Development in Higher Education*, (1).

Bakhtin, M. M., & Holquist, M. (1981). *The dialogic imagination: Four essays*. University of Texas Press.

Bakhtin, M. M., Holquist, M., McGee, V., & Emerson, C. (1986). *Speech genres and other late essays*. University of Texas Press.

Balan, A., & Jönsson, A. (2018). Increased explicitness of assessment criteria: Effects on student motivation and performance. *Frontiers in Education*, *3*, 81. doi:10.3389/feduc.2018.00081

Bandura, A. (1977). Self-efficacy: Toward a unifying theory of behavioral change. *Psychological Review*, *84*(2), 191–215. doi:10.1037/0033-295X.84.2.191 PMID:847061

Compilation of References

Barba-Martín, R. A., Bores-García, D., Hortigüela-Alcalá, D., & González-Calvo, G. (2020). Evaluación formativa con los estudiantes en prácticas para reducir la brecha teoría-práctica en la formación inicial del profesorado. *Educacion Fisica y Deporte*, *39*(1). doi:10.17533/udea.efyd.v39n1a02

Barba-Martín, R. A., Hortigüela-Alcalá, D., & Pérez-Pueyo, A. (2020). Evaluar en educación física: Análisis de las tensiones existentes y justificación del empleo de la evaluación formativa y compartida. *Educacion Fisica y Deporte*, *39*(1). doi:10.17533/udea.efyd.v39n1a03

Barba-Martín, R. A., & Hotigüela-Alcalá, D. (2022). Si la evaluación es aprendizaje, he de formar parte de la misma. Razones que justifican la implicación del alumnado. *Revista Iberoamericana de Evaluación Educativa*, *15*(1), 9–22. doi:10.15366/riee2022.15.1.001

Baron, R. A. (2006). Opportunity recognition as pattern recognition: How entrepreneurs "connect the dots" to identify new business opportunities. *The Academy of Management Perspectives*, *20*(1), 104–119. doi:10.5465/amp.2006.19873412

Barr, R. B., & Tagg, J. (1995). "From Teaching to Learning—A New Paradigm For Undergraduate Education." Change. *Change*, *27*(6), 12–26. doi:10.1080/00091383.1995.10544672

Bayazıt, İ. (2011). Prospective teachers' understanding of graphs. *Gaziantep University Journal of Social Sciences*, *10*(4), 1325–1346.

Bearman, M., & Ajjawi, R. (2018). From "seeing through" to "seeing with": Assessment criteria and the myths of transparency. *Frontiers in Education*, *3*(96), 1–8. doi:10.3389/feduc.2018.00096

Bearman, M., & Ajjawi, R. (2021). Can a rubric do more than be transparent? Invitation as a new metaphor for assessment criteria. *Studies in Higher Education*, *46*(2), 359–368. doi:10.1080/03075079.2019.1637842

Becker, A. (2016). Student-generated scoring rubrics: Examining their formative value for improving ESL students' writing performance. *Assessing Writing*, *29*, 15–24. doi:10.1016/j.asw.2016.05.002

Becker, G. S. (1993). *Human Capital*. University of Chicago Press. doi:10.7208/chicago/9780226041223.001.0001

Becker, G. S. (2011). Foreword. In *The Oxford Handbook of Human Capital, edited by A. BurtonJones and J.-C. Spender, ppxiii–ppxxvi*. Oxford University Press. doi:10.2307/j.ctt6wq7jd.3

Beichner, R. (1994). Testing student interpretation of kinematics graphs. *American Journal of Physics*, *62*(8), 750–762. doi:10.1119/1.17449

Bektaslı, B., & White, A. L. (2012). The relationships between logical thinking, gender, and kinematics graph interpretation skills. *Eurasian Journal of Educational Research*, *48*, 1–20.

Bell, A., Mladenovic, R., & Price, M. (2013). Students' perceptions of the usefulness of marking guides, grade descriptors and annotated exemplars. *Assessment & Evaluation in Higher Education*, *38*(7), 769–788. doi:10.1080/02602938.2012.714738

Beltrán-Carrillo, V. J., & Devís-Devís, J. (2019). El pensamiento del alumnado inactivo sobre sus experiencias negativas en educación física: Los discursos del rendimiento, salutismo y masculinidad hegemónica. *RICYDE. Revista Internacional de Ciencias del Deporte*, *55*(15), 20–34. doi:10.5232/ricyde2019.05502

Benesch, S. (2001). *Critical English for Academic Purposes: theory, politics and practice*. Routledge. doi:10.4324/9781410601803

BERA-RSA. (2014). *The role of research in teacher education. Reviewing the evidence. Interim report of the BERA-RSA inquiry*. London: BERA. https://www.bera.ac.uk/wpcontent/uploads/2014/02/BERA-RSA-Interim-Report.pdf

Bergevoet, R. H., & Woerkum, C. V. (2006). Improving the entrepreneurial competencies of Dutch dairy farmers through the use of study groups. *Journal of Agricultural Education and Extension*, *12*(1), 25–39. doi:10.1080/13892240600740852

Berg, M. A., & Huang, J. (2015). Improving in-service teachers' effectiveness: K-12 academic literacy for the linguistically diverse. *Functional Linguistics*, *2*(1), 1–21. doi:10.118640554-015-0017-6

Bernstein, B. (1996). *Pedagogy, symbolic control and identity: Theory, research and critique*. Taylor and Francis.

Beverland, M., & Lindgreen, A. (2010). What Makes a Good Case Study? A Positivist Review of Qualitative Case Research Published in Industrial Marketing Management, 1971–2006. *Industrial Marketing Management*, *39*(1), 56–63. doi:10.1016/j.indmarman.2008.09.005

Bharuthram, S. (2012). Making a case for the teaching of reading across the curriculum in higher education. *South African Journal of Education*, *32*(2), 205–214. doi:10.15700aje.v32n2a557

Biggs, J. (1999). *Teaching for quality learning at university: what the student does*. Open University Press.

Biggs, J. (2005). *Calidad del aprendizaje universitario*. Narcea.

Biggs, J., & Tang, C. (2011). *Teaching for Quality Learning at University*. Open University Press.

Bird, B. (1995). Towards a Theory of Entrepreneurial Competency. *Advances in Entrepreneurship, Firm, Emergence, and Growth*, *2*(1), 51–72. doi:10.1108/S1074-754020190000021011

Black, P., & Wiliam, D. (2003). 'In Praise of Educational Research': Formative assessment. *British Educational Research Journal*, *29*(5), 623–637. doi:10.1080/0141192032000133721

Blaz, D. (2001). *A collection of performance tasks and rubrics: Foreign languages*. Eye On Education.

Blenker, P., Elmholdt, S. T., Frederiksen, S. H., Korsgaard, S., & Wagner, K. (2014). Methods in entrepreneurship education research: A review and integrative framework. *Education + Training*, *56*(8/9), 697–715. doi:10.1108/ET-06-2014-0066

Bloxham, S., & Boyd, P. (2007). *Developing effective assessment in higher education: a practical guide: a practical guide*. McGraw-Hill Education.

Bloxham, S., Boyd, P., & Orr, S. (2011). Mark my words: The role of assessment criteria in UK higher education grading practices. *Studies in Higher Education*, *36*(6), 655–670. doi:10.1080/03075071003777716

Bloxham, S., & Campbell, L. (2010). Generating dialogue in assessment feedback: Exploring the use of interactive cover sheets. *Assessment & Evaluation in Higher Education*, *35*(3), 291–300. doi:10.1080/02602931003650045

Bohm, D. (1996). *On dialogue*. Routledge.

Bolliger, D. U., & Halupa, C. (2018). Online student perceptions of engagement, transactional distance, and outcomes. *Distance Education*, *39*(3), 299–316. doi:10.1080/01587919.2018.1476845

Bolliger, D. U., & Martin, F. (2018). Instructor and student perceptions of online student engagement strategies. *Distance Education*, *39*(4), 568–583. doi:10.1080/01587919.2018.1520041

Bolton, F. C. (2006). Rubrics and adult learners: Andragogy and assessment. *Assessment Update*, *18*(3), 5–6.

Bonanno, A. (2017). *The Legitimation Crisis of Neoliberalism: The State, Will-Formation, and Resistance*. Springer. doi:10.1057/978-1-137-59246-0

Bond, B. (2020). *Making language visible in the university*. Multilingual Matters.

Boote, S. K. (2014). Assessing and understanding line graph interpretations using a scoring rubric of organized cited factors. *Journal of Science Teacher Education*, *25*(3), 333–354. doi:10.100710972-012-9318-8

Bordin, E. S. (1983). A working alliance based model of supervision. *The Counseling Psychologist*, *11*(1), 35–41. doi:10.1177/0011000083111007

Bores-García, D., Hortigüela-Alcalá, D., González-Calvo, G., & Barba-Martín, R. A. (2020). Peer Assessment in Physical Education: A Systematic Review of the Last Five Years. *Sustainability (Basel)*, *12*(21), 9233. doi:10.3390u12219233

Bores-García, D., Hortigüela-Alcalá, D., Hernando-Garijo, A., & González-Calvo, G. (2021). Analysis of student motivation towards body expression through the use of formative and share assessment. *Retos*, *40*(40), 198–208. doi:10.47197/retos.v1i40.83025

Boud, D. (2010). *Seven Propositions for Assessment Reform in Perspective*. Routledge.

Boud, D. (2013). *Enhancing learning through self-assessment*. Routledge. doi:10.4324/9781315041520

Boud, D., & Falchikov, N. (2006). Aligning Assessment with Long-Term Learning. *Assessment & Evaluation in Higher Education*, *31*(4), 399–413. doi:10.1080/02602930600679050

Bourgeois, A., Balcon, M. P., & Riiheläinen, J. M. (2016). *Entrepreneurship Education at School in Europe. Eurydice 2016 Report.* Education, Audiovisual and Culture Executive Agency, European Commission. https://eacea.ec.europa.eu/national-policies/Bourgeois

Bourke, J. (2005). The grammar we teach. *Reflections on English Language Teaching*, *4*(2), 85–97.

Bovill, C., & Bulley, C. J. (2011). A model of active student participation in curriculum design: exploring desirability and possibility. In C. Rust (Ed.), *Improving Student Learning (18) Global theories and local practices: Institutional, disciplinary and cultural variations* (pp. 176–188). The Oxford Centre for Staff and Learning Development.

Bowles, M. (2010). *The think-aloud controversy*. Routledge.

Boyd, N. G., & Vozikis, G. S. (1994). The influence of self-efficacy on the development of entrepreneurial intentions and actions. *Entrepreneurship Theory and Practice*, *18*(4), 63–77. doi:10.1177/104225879401800404

Bradley, E., Anderson, S., & Eagle, L. A. (2020). Use of a marking rubric and self-assessment to provide feedforward to level 5 undergraduate Sport students: student perceptions, performance and marking efficiency. *Journal of Learning Development in Higher Education, 18.*

Bradley, L., Lindström, B., & Rystedt, H. (2010). Rationalities of collaboration for language learning in a wiki. *ReCALL*, *22*(2), 247–265. doi:10.1017/S0958344010000108

Brennan, R., & Vos, L. (2021). Introduction to teaching marketing. In R. Brennan & L. Vos (Eds.), *Teaching marketing*. Edward Elgar Publishing. doi:10.4337/9781789907896.00008

Bresciani, M. J., Zelna, C. L., & Anderson, J. A. (2004). Criteria and rubrics. Assessing Student Learning and Development: A Handbook for Practitioners, National Association of Student Personnel Administrators, 29–37.

Bresciani, M. J., Zelna, C. L., & Anderson, J. A. (2004). Criteria and rubrics. In *Assessing Student Learning and Development: A Handbook for Practitioners* (pp. 29–37). National Association of Student Personnel Administrators.

Bridge, S., O'Neill, K., & Cromie, S. (2003). *Understanding enterprise. Entrepreneurship and Small Business*. Palgrave McMillan.

Brisk, M., & Zisselsberger, M. (2011). "We've let them in on a secret": Using SFL theory to improve the teaching of writing to bilingual learners. In T. Lucas (Ed.), *Teacher preparation for linguistically diverse classrooms: A resource for teacher educators* (pp. 111–126). Routledge.

Brookhart, S. M. (2013). *How to create and use rubrics for formative assessment and grading*. ASCD.

Brookhart, S. M. (2018). Appropriate criteria: Key to effective rubrics. *Frontiers in Education*, *3*(22), 1–12.

Brookhart, S. M. (2018). Appropriate Criteria: Key to Effective Rubrics. *Frontiers in Education*, *3*, 22. https://www.frontiersin.org/articles/10.3389/feduc.2018.00022. doi:10.3389/feduc.2018.00022

Brookhart, S. M., & Chen, F. (2015). The quality and effectiveness of descriptive rubrics. *Educational Review*, *67*(3), 343–368. doi:10.1080/00131911.2014.929565

Brook, J., Catlin, S., DeLuca, C., Doe, C., Huntly, A., & Searle, M. (2010). Conceptions of doctoral education: The PhD as pathmaking. *Reflective Practice*, *11*(5), 657–668. doi:10.1080/14623943.2010.516981

Brophy, J. (1985). Teachers' expectations, motives, and goals for working with problem students. In C. Ames & R. Ames (Eds.), Research on motivation in education: Vol. 2. The classroom milieu (pp. 175–213). New York: Academic Press.

Brophy, M. and Kiely, T. (2002). Competencies: a new sector. *Journal of European Industrial Training, 26* (2/3/4), 165-176. doi:10.1108/03090590210422049

Brown, G. A., Bakhtar, M., & Youngman, M. B. (1984). Toward a typology of lecturing styles. *The British Journal of Educational Psychology*, *54*(1), 93–100. doi:10.1111/j.2044-8279.1984.tb00848.x

Brown, G. T. L., & Harris, L. R. (2013). Student self-assessment. In J. H. McMillan (Ed.), *The SAGE handbook of research on classroom assessment* (pp. 367–393). doi:10.4135/9781452218649.n21

Brown, S., & Glasner, A. (2003). *Evaluar en la Universidad. Problemas y nuevos enfoques*. Narcea.

Brown, S., & Pickford, R. (2013). *Assessing Skill and Practice*. Routledge.

Buber, M. (1947). *Between man and man* (R. G. Smith, Trans.). Routledge & Kegan Paul.

Buckley, A. (2021). Crisis? What crisis? Interpreting student feedback on assessment. *Assessment & Evaluation in Higher Education*, *46*(7), 1008–1019. doi:10.1080/02602938.2020.1846015

Bui, H. T. (2014). Student–supervisor expectations in the doctoral supervision process for business and management students. *Business and Management Education in HE*, *1*(1), 12–27. doi:10.11120/bmhe.2014.00006

Butler, D. L., & Winne, P. H. (1995). Feedback and self-regulated learning: A theoretical synthesis. *Review of Educational Research*, *65*(3), 245–281. doi:10.3102/00346543065003245

Butz, D., & Besio, K. (2009). Autoethnography. *Geography Compass*, *3*(5), 1660–1674. doi:10.1111/j.1749-8198.2009.00279.x

Calatayud, M. A. (2007). La evaluación como instrumento de aprendizaje y mejora. Una luz al fondo. In A. M. Calatayud (Coord). La evaluación como instrumento de aprendizaje. Estrategias y técnicas. (p. 9 – 23). Madrid.

Calatayud, M. A. (2008). Establecer la Cultura de la Autoevaluación. *Padres y Maestros*, *314*(febrero), 30–34.

Calatayud, M. A. (2018). La autoevaluación. Una propuesta formativa e inovadora. *Revista Ibero-Americana de Educação*, *72*(2), 135–152. doi:10.35362/rie7623081

Calderon, M. E., & Slakk, S. (2018). *Teaching reading to English learners, grades 6-12: A framework for improving achievement in the content areas*. Corwin Press.

Camargo, S., Chang, A., Maeda, Y. (2023, April). *Effect of rubrics on motivation and performance: Meta-analysis study*. [Paper Presentation] American Educational Research Association.

Campbell, A. (2005). Application of ICT and rubrics to the assessment process where professional judgement is involved: The features of an e-marking tool. *Assessment & Evaluation in Higher Education*, *30*(5), 529–537. doi:10.1080/02602930500187055

Carless, D., & Boud, D. (2018). The development of student feedback literacy: Enabling uptake of feedback. *Assessment & Evaluation in Higher Education*, *43*(8), 1315–1325. doi:10.1080/02602938.2018.1463354

Carless, D., & Chan, K. (2017). Managing dialogic use of exemplars. *Assessment & Evaluation in Higher Education*, *42*(6), 930–941. doi:10.1080/02602938.2016.1211246

Carroll, N., & Conboy, K. (2020). Normalising the "new normal": Changing tech-driven work practices under pandemic time pressure. *International Journal of Information Management*, *55*, 102186. doi:10.1016/j.ijinfomgt.2020.102186 PMID:32836643

Castañeda, L., Talborn Björkvi, S., Tilly, A., Minin, D., Hernández, I., & Hämäläinen, M. (2021). *Aprendizaje conectado como práctica sistémica para procesos de desarrollo profesional docente: un estudio de caso basado en la combinación de estrategias*.

Çelik, D., & Sağlam Arslan, A. (2012). The analysis of teacher candidates' translating skills in multiple representations. *Elementary Education Online*, *11*(1), 239–250.

Chaker, R., Bouchet, F., & Bachelet, R. (2022). How do online learning intentions lead to learning outcomes? The mediating effect of the autotelic dimension of flow in a MOOC. *Computers in Human Behavior*, *134*, 107306. doi:10.1016/j.chb.2022.107306

Chan, C. K. Y., & Lee, K. K. W. (2021). Reflection literacy: A multilevel perspective on the challenges of using reflections in higher education through a comprehensive literature review. *Educational Research Review*, *32*, 100376. doi:10.1016/j.edurev.2020.100376

Chan, C. K. Y., & Luo, J. (2021a). Exploring teacher perceptions of different types of 'feedback practices' in higher education: Implications for teacher feedback literacy. *Assessment & Evaluation in Higher Education*, *47*(1), 61–76. doi:10.1080/02602938.2021.1888074

Chan, C. K. Y., & Luo, J. (2021b). A four-dimensional conceptual framework for student assessment literacy in holistic competency development. *Assessment & Evaluation in Higher Education*, *46*(3), 451–466. doi:10.1080/02602938.2020.1777388

Chan, Z., & Ho, S. (2019). Good and bad practices in rubrics: The perspectives of students and educators. *Assessment & Evaluation in Higher Education*, *44*(4), 533–545. doi:10.1080/02602938.2018.1522528

Chapman, V. G., & Inman, M. D. (2009). A conundrum: Rubrics or creativity/metacognitive development? *Educational Horizons*, *87*(3), 198–202.

Chappuis, J., Stiggins, R. J., Chappuis, S., & Arter, J. (2012). *Classroom assessment for student learning: Doing it right-using it well*. Pearson.

Chasteen, S. V., Pepper, R. E., Caballero, M. D., Pollock, S. J., & Perkins, K. K. (2012). Colorado upper-division electrostatics diagnostic: A conceptual assessment for the junior level. *Physical Review Special Topics. Physics Education Research*, *8*(2), 020180. doi:10.1103/PhysRevSTPER.8.020108

Chasteen, S. V., & Scherr, R. E. (2020). Developing the physics teacher education program analysis rubric: Measuring features of thriving programs. *Physical Review. Physics Education Research*, *16*(1), 010115. doi:10.1103/PhysRevPhysEducRes.16.010115

Chen, F., & Andrade, H. (2016). The impact of criteria-referenced formative assessment on fifth grade students' theater arts achievement. *The Journal of Educational Research*, *109*, 1–10.

Chen, F., Lui, A., Andrade, H., Valle, C., & Mir, H. (2017). Criteria-referenced formative assessment in the arts. *Educational Assessment, Evaluation and Accountability*, *29*(3), 297–314. doi:10.100711092-017-9259-z

Cheng, M. W. T., & Chan, C. K. Y. (2019). *An experimental test: Using rubrics for reflective writing to develop reflection*. Elsevier. doi:10.1016/j.stueduc.2019.04.001

Chng, L. S., & Lund, J. (2018). Assessment for Learning in Physical Education: The What, Why and How. *Journal of Physical Education, Recreation & Dance*, *89*(8), 29–34. doi:10.1080/07303084.2018.1503119

Chowdhury, F. (2018). Application of Rubrics in the Classroom: A Vital Tool for Improvement in Assessment, Feedback and Learning. *International Education Studies*, *12*(1), 61. doi:10.5539/ies.v12n1p61

Christie, M., Grainger, P. R., Dahlgren, R., Call, K., Heck, D., & Simon, S. E. (2015). Improving the quality of assessment grading tools in master of education courses: A comparative case study in the scholarship of teaching and learning. *The Journal of Scholarship of Teaching and Learning*, *15*(5), 22–35. doi:10.14434/josotl.v15i5.13783

Chugh, R., Macht, S., & Harreveld, B. (2021). Supervisory Feedback to Postgraduate Research Students: A Literature Review. *Assessment & Evaluation in Higher Education*, 1–15. doi:10.1080/02602938.2021.1955241

Cifrian, E., Andrés, A., Galán, B., & Viguri, J. R. (2020). Integration of different assessment approaches: Application to a project-based learning engineering course. *Education for Chemical Engineers*, *31*, 62–75. doi:10.1016/j.ece.2020.04.006

Cimpian, A., Arce, H.-M. C., Markman, E. M., & Dweck, C. S. (2007). Subtle linguistic cues affect children's motivation. *Psychological Science*, *18*(4), 314–316. doi:10.1111/j.1467-9280.2007.01896.x PMID:17470255

Clouder, L., Karakus, M., Cinotti, A., Ferreyra, M. V., Fierros, G. A., & Rojo, P. (2020). Neurodiversity in higher education: A narrative synthesis. *Higher Education*, *80*(4), 757–778. doi:10.100710734-020-00513-6

Cochran-Smith, M., & Lytle, S. L. (1992). Communities for teacher research: Fringe or forefront? *American Journal of Education*, *100*(3), 298–324. doi:10.1086/444019

Cochran-Smith, M., & Lytle, S. L. (1993). *Inside/outside: Teacher research and knowledge*. Teachers College Press.

Cochran-Smith, M., & Lytle, S. L. (1999). Relationships of knowledge and practice: Teacher learning in communities. *Review of Research in Education*, *24*, 249–305.

Cochran-Smith, M., & Lytle, S. L. (2009). *Inquiry as Stance: Practitioner Research for the Next Generation*. Teachers College Press.

Cochran-Smith, M., & Zeichner, K. M. (2005). *Studying teacher education: The report of the AERA panel on research and teacher education*. American Educational Research Association.

Cockett, A., & Jackson, C. (2018). The use of assessment rubrics to enhance feedback in higher education: An integrative literature review. *Nurse Education Today, 69*, 8–13. doi:10.1016/j.nedt.2018.06.022

Cole, R., Lantz, J., Ruder, S., Reynders, G., & Stanford, C. (2018). Enhancing learning by assessing more than content knowledge. ASEE Annual Conference & Exposition, Salt Lake City, Utah, Covill, A. E. (2012). College students' use of a writing rubric: Effect on quality of writing, self-efficacy, and writing practices. *Journal of Writing Assessment*, *5*(1).

Cook-Sather, A., Bovill, C., & Felten, P. (2014). *Engaging students as partners in teaching and learning: A guide for faculty*. Jossey-Bass.

Coştu, F., Beler, Ş., & Coştu, B. (2017). Students' graphic skills and their difficulties about photosynthesis topic. *Scientific Educational Studies,* 41-63.

Cotoi, E., Bodoasca, T., Catana, L., & Cotoi, I. (2011). Entrepreneurship European development strategy in the field of education. *Procedia: Social and Behavioral Sciences*, *15*, 3490–3494. doi:10.1016/j.sbspro.2011.04.323

Cotton, A. H. (2001). Private thoughts in public spheres: Issues in reflection and reflective practices in nursing. *Journal of Advanced Nursing*, *36*(4), 512–519. doi:10.1046/j.1365-2648.2001.02003.x PMID:11703545

Coulehan, J., & Granek, I. A. (2012). Commentary: "I Hope I'll continue to grow": Rubrics and reflective writing in medical education. *Academic Medicine*, *87*(1), 8–10. doi:10.1097/ACM.0b013e31823a98ba PMID:22201632

Coxhead, A. (2000). A new academic word list. *TESOL Quarterly*, *34*(2), 213–238. doi:10.2307/3587951

Cox, K., Imrie, B. W., & Miller, A. (2014). *Student assessment in higher education: A handbook for assessing performance*. Routledge. doi:10.4324/9781315042107

Creswell, J. W. (1998). *Qualitative inquiry and research design: Choosing among five traditions*. Sage.

Creswell, J. W., & Creswell, J. D. (2023). *Research design: qualitative, quantitative, and mixed methods approaches* (6th ed.). SAGE.

Creswell, J., & Creswell, J. D. (2018). *Research Design: Qualitative, Quantitative, and Mixed Methods Approaches* (5th ed.). Sage Publications.

Cuban, L. (1993). The lure of curriculum reform and its pitiful history. *Phi Delta Kappan*, *75*(2), 181–185.

Cuenca, L., Alarcón Valero, F., Boza, A., Fernández Diego, M., Ruiz Font, L., Gordo Monzó, M. L., & Alemany Díaz, M. D. M. (2016). Rubric to Assess the Competence OF Innovatoon, Creativity and Entrepreneurship in Bachelor Degree. *Brazilian Journal of Operations & Production Management*, *13*(1), 118–123. doi:10.14488/BJOPM.2016.v13.n1.a14

Cullen, R., & Harris, M. (2009). Assessing Learner-centredness through Course Syllabi. *Assessment & Evaluation in Higher Education*, *34*(1), 115–125. doi:10.1080/02602930801956018

Darling-Hammond, L. (2016). Research on teaching and teacher education and its influences on policy and practice. *Educational Researcher*, *45*(2), 83–91. doi:10.3102/0013189X16639597

Darling-Hammond, L. (2017). Teacher education around the world: What can we learn from international practice? *European Journal of Teacher Education*, *40*(3), 291–30. doi:10.1080/02619768.2017.1315399

Dawson, P. (2017). Assessment rubrics: Towards clearer and more replicable design, research and practice. *Assessment & Evaluation in Higher Education*, *42*(3), 347–360. doi:10.1080/02602938.2015.1111294

de Boer, I., de Vegt, F., Pluk, H., & Latijnhouwers, M. (2021). *Rubrics-a Tool for Feedback and Assessment Viewed from Different Perspectives: Enhancing Learning and Assessment Quality*. Springer. doi:10.1007/978-3-030-86848-2

De Kleijn, R. A., Meijer, P. C., Brekelmans, M., & Pilot, A. (2015). Adaptive research supervision: exploring expert thesis supervisors' practical knowledge. *Higher Education Research & Development, 34*(1), pp.117-130.

De La Croix, A., & Veen, M. (2018). EYE-OPENER The reflective zombie: Problematizing the conceptual framework of reflection in medical education. *Perspectives on Medical Education*, 7(6), 394–400. doi:10.1007/S40037-018-0479-9 PMID:30353284

de Lourdes Cárcamo-Solís, M., del Pilar Arroyo-López, M., del Carmen Alvarez-Castañón, L., & García-López, E. (2017). Developing entrepreneurship in primary schools. The Mexican experience of "My first enterprise: Entrepreneurship by playing". *Teaching and Teacher Education*, *64*, 291–304. doi:10.1016/j.tate.2017.02.013

de Oliveira, L., Smith, S. L., Axelrod, D., Diaz, E., & Vicentini, C. (2021). Supporting Academic Language Development for Multilingual Learners Across Content Areas Through the Identification of Textual Features. *Journal of Narrative and Language Studies*, *9*(17), 227–242.

Deci, E. L., & Ryan, R. M. (1985). *Intrinsic motivation and self-determination in human behavior*. Plenum Press. doi:10.1007/978-1-4899-2271-7

Decreto-lei n.º 43/2007 de 22 de fevereiro. Diário da República n.º 38/2007 - Série I de 2007-02-22, 1320-1328. Lisboa: Ministério da Educação. https://data.dre.pt/eli/dec-lei/43/2007/02/22/p/dre/pt/html

Decreto-lei n.º 74/2006 de 24 de março. Diário da República n.º 60/2006 - Série I-A de 2006-03-24, 2242-2257. Lisboa: Ministério da Ciência, Tecnologia e Ensino Superior. https://data.dre.pt/eli/dec-lei/74/2006/03/24/p/dre/pt/html

Decreto-lei n.º 79/2014 de 14 de maio. Diário da República n.º 92/2014 - Série I de 2014-05-14, 2819-2828. Lisboa: Ministério da Educação e Ciência. https://data.dre.pt/eli/dec-lei/79/2014/05/14/p/dre/pt/html

Del Rio-Roberts, M. (2011). How I Learned to Conduct Focus Groups. *Qualitative Report*, *16*(1), 312–315.

Demirci, N., & Uyanık, F. (2009). The correlation between tenth grade students' understanding and interpreting graphs and their kinematics achievment. *Necatibey Faculty of Education. Electronic Journal of Science and Mathematics Education*, *3*(2), 22–51.

Devlin, M. J., Mutnick, A., Balmer, D., & Richards, B. F. (2010). Clerkship-based reflective writing: A rubric for feedback. *Medical Education*, *44*(11), 1143–1144. doi:10.1111/j.1365-2923.2010.03815.x PMID:20946509

Dewey, J. (2006). How we think. In *How we think*. D C Heath. doi:10.1037/10903-000

Dochy, F., Segers, M., & Dierick, S. (2002). Nuevas Vías de Aprendizaje y Enseñanza y sus Consequencias: una Nueva Era de Evaluación. *Revista de Docencia Universitária, 2*(2) Education: Classroom Teachers' Perspective. *The Journal of Educational Research*, *101*(2), 99–108. doi:10.3200/JOER.101.2.99-112

Docktor, J. L., Dornfeld, J., Frodermann, E., Heller, K., Hsu, L., Jackson, K. A., Mason, A., Ryan, Q. X., & Yang, J. (2016). Assessing student written problem solutions: A problem-solving rubric with application to introductory physics. *Physical Review. Physics Education Research*, *12*(1), 010130. doi:10.1103/PhysRevPhysEducRes.12.010130

Donaghy, M. E., & Morss, K. (2009). *Guided reflection: A framework to facilitate and assess reflective practice within the discipline of physiotherapy*. Taylor and Francis. doi:10.1080/095939800307566

Dornyei, Z. (2007). *Research methods in applied linguistics: quantitative qualitative, and mixed methodologies*. Oxford University Press.

Doyle, W. (2006). Ecological management and classroom management. In *Handbook of Classroom Management*. Lawrence Erlbaum.

Driscoll, J., & Teh, B. (2001). The potential of reflective practice to develop individual orthopaedic nurse practitioners and their practice. *Journal of Orthopaedic Nursing*, *5*(2), 95–103. doi:10.1054/joon.2001.0150

Dunne, J., & Ryan, S. M. (2016). Enhancing Professional Development and Supporting Students on Work-Placement by Peer-Peer Learning Using an Online Reflective Blog Assessment. *Irish Journal of Academic Practice*, *5*(1), 1. doi:10.21427/D7HT51

Dweck, C. S. (1999). *Self-theories: their role in motivation, personality, and development*. Psychology Press.

Dwyer, C. A., Millett, C. M., & Payne, D. G. (2006) A Culture of Evidence: Postsecondary Assessment and Learning Outcomes. Recommendations to Policymakers and the Higher Education Community. *Educational Testing Service*. https://files.eric.ed.gov/fulltext/ED500004.pdf

Dysthe, O. (2002). Professors as mediators of academic text cultures: An interview study with advisors and master's degree students in three disciplines in a Norwegian university. *Written Communication*, *19*(4), 493–544. doi:10.1177/074108802238010

Ecclestone, K. (2001). 'I know a 2: 1 when I see it': Understanding criteria for degree classifications in franchised university programmes. *Journal of Further and Higher Education*, *25*(3), 301–313. doi:10.1080/03098770126527

Edwards, F. (2020). Engaging tertiary educators in the development of their assessment literacy. *Teachers and Curriculum*, *20*(1), 87–92. doi:10.15663/tandc.v20i1.345

Ellis, C., Adams, T. E., & Bochner, A. P. (2011). Autoethnography: An overview. Historical Social Research, 36(4 Elliot, J., (1991). Action research for educational change. OUP: Buckingham

Empson, W. (1930). *Seven types of ambiguity*. Chatto and Windus.

Empson, W. (1951). *The Structure of Complex Words*. Chatto & Windus.

Ene, E., & Kosobucki, V. (2016). Rubrics and corrective feedback in ESL writing: A longitudinal case study of an L2 writer. *Assessing Writing*, *30*, 3–20. doi:10.1016/j.asw.2016.06.003

Ene, E., & Upton, T. (2014). Learner uptake of teacher electronic feedback in ESL composition. *System*, *46*, 80–95. doi:10.1016/j.system.2014.07.011

English, F. (1978). *Quality control in curriculum development*. American Association of School Administrators.

EQF. (2005). *Descriptors defining levels in the European Qualifications Framework*. EQF. https://ec.europa.eu/ploteus/tr/node/1440

Ercan, O., Coştu, F., & Coştu, B. (2018). Identifying pre-service science teachers' difficulties about graph drawings. *Kastamonu Education Journal*, *26*(6), 1929–1938.

Ergül, N. R. (2018). Pre-service science teachers' construction and interpretation of graphs. *Universal Journal of Educational Research*, *6*(1), 139–144. doi:10.13189/ujer.2018.060113

Ericsson, K., & Simon, H. (1993). *Protocol analysis: verbal reports as data* (2nd ed.). MIT Press. doi:10.7551/mitpress/5657.001.0001

Eryılmaz Toksoy, S. (2020). Investigation of 11th grade students' knowledge about explanation of motion types and drawing, interpreting related graphs. *Bolu Abant İzzet Baysal University Journal of Faculty of Education*, *20*(3), 1423–1441.

Etkina, E., Heuvelen, A. V., White-Brahmia, S., Brookes, D. T., Gentile, M., Murthy, S., Rosengrant, D., & Warren, A. (2006). Scientific abilities and their assessment. *Physical Review Special Topics. Physics Education Research*, *2*(2), 020103. doi:10.1103/PhysRevSTPER.2.020103

European Commission. (2012). *Vocational education and training for better skills, growth and jobs*. EC. https://eur-lex.europa.eu/legal-content/EN/TXT/PDF/?uri=CELEX:52012SC0375&from=EN

European Commission. (2013). *Entrepreneurship 2020 action plan: Reigniting the entrepreneurial spirit in Europe*. European Commission.

Faletič, S., & Planinšič, G. (2020). How the introduction of self-assessment rubrics helped students and teachers in a project laboratory course. *Physical Review. Physics Education Research*, *16*(2), 020136. doi:10.1103/PhysRevPhysEducRes.16.020136

Fang, Z., & Wang, Z. (2011). Beyond rubrics: Using functional language analysis to evaluate student writing. *Australian Journal of Language and Literacy*, *34*(2), 147–165. doi:10.1007/BF03651853

Farquhar, J. D. (2012). *Case Study Research for Business*. SAGE. doi:10.4135/9781446287910

Farrell, O., & Brunton, J. (2020). A balancing act: A window into online student engagement experiences. *International Journal of Educational Technology in Higher Education*, *17*(1), 25. doi:10.118641239-020-00199-x

Fastré, G., & Van Gils, A. (2007). Competence development in Entrepreneurship. The Role of University Education. In M.K. McCuddy (pp. 385-398) The challenges of educating people to lead in a challenging world. Springer, Dordrecht. doi:10.1007/978-1-4020-5612-3_19

Fecho, B. (2011). *Teaching for the students: Habits of heart, mind, and practice in the engaged classroom*. Teachers College Press.

Fernandes, D. (2021). *Para uma Fundamentação e Melhoria das Práticas de Avaliação Pedagógica. Projeto de Monitorização Acompanhamento e Investigação em Avaliação pedagógica*. Universidade de Lisboa - Instituto de Educação, 3.

Fernández-Balboa, J. M. (2005). La autoevaluación como práctica promotora de la democracia y la dignidad. En A. Sicilia, & J. M. Fernández-Balboa (Coords.), La otra cara de la educación física: la educación física desde una perspectiva crítica (pp. 127-158). Inde.

Fernández-Balboa, J. M. (2007). Dignity and democracy in the college classroom: The practice of self-evaluation. In R. A. Goldstein (Ed.), *Useful Theory: Making Critical Education Practical* (pp. 105–128). Peter Lang Publishing.

Ferreras-Garcia, R., Hernández-Lara, A. B., & Serradell-López, E. (2019). Entrepreneurial competences in a higher education business plan course. *Education + Training*, *61*(7/8), 850–869. doi:10.1108/ET-04-2018-0090

Fiet, J. (2001). The Theoretical Side of Teaching Entrepreneurship. *Journal of Business Venturing*, *16*(1), 1–24. doi:10.1016/S0883-9026(99)00041-5

Figueiredo-Nery, M., N., A., & Figueiredo, P. (2008). Forming entrepreneurial mindsets? Preliminary evidence of teaching practices from primary schools in a developing area in south America. *Journal of Technology Management & Innovation*, *3*(2), 1–17.

Finnish Ministry of Education and Culture. (2016). *Finnish education in a nutshell*. FMEC. https://www.oph.fi/en/statistics-and-publications/publications/finnish-education-nutshell

Finnish Ministry of Education. (2009). *Guidelines for entrepreneurship education*. FME. https://julkaisut.valtioneuvosto.fi/handle/10024/78871

Finnish National Agency for Education. (2016). National Core Curriculum for Basic Education. [Helsinki.]. *Publications*, *2016*, 5.

Fitt, M. H., Walker, A. E., & Leary, H. M. (2009). Assessing the quality of doctoral dissertation literature reviews in instructional technology. *Paper presented at the Annual meeting of the American Educational Research Association San Diego*, 3-15.

Fives, H., & Buehl, M. M. (2012). Spring cleaning for the "messy" construct of teachers' beliefs: What are they? Which have been examined? What can they tell us. *Apa Educational Psychology Handbook*, *2*, 471–499. doi:10.1037/13274-019

Flavell, J. H. (1979). Metacognition and cognitive monitoring: A new area of cognitive–developmental inquiry. *The American Psychologist*, *34*(10), 906–911. doi:10.1037/0003-066X.34.10.906

Fletcher, R. B., Meyer, L. H., Anderson, H., Johnston, P., & Rees, M. (2012). Faculty and students conceptions of assessment in higher education. *Higher Education*, *64*(1), 119–133. doi:10.100710734-011-9484-1

Flexner, A. (1910). Medical education in The United States and Canada a report to the Carnegie Foundation for the advancement of teaching. Bulletin Number Four. Boston: The Merrymount Press (Reproduces in 1972).

Flick, U. (2007). Managing quality in qualitative research. *Sage (Atlanta, Ga.)*.

Flores, M. A. (2020). Feeling like a student but thinking like a teacher: A study of the development of professional identity in initial teacher education. *Journal of Education for Teaching*, *46*(2), 145–158. doi:10.1080/02607476.2020.1724659

Fontaine, L. M. (2013). Introduction: Choice in contemporary systemic functional theory. In L. Fontaine, Tom Bartlett, & G. O'Grady (Eds.), Systemic functional linguistics: Exploring choice. (pp. 1-12). Cambridge University Press.

Fraile, J., Panadero, E., & Pardo, R. (2017). Co-creating rubrics: The effects on self-regulated learning, self-efficacy and performance of establishing assessment criteria with students. *Studies in Educational Evaluation*, *53*, 69–76. doi:10.1016/j.stueduc.2017.03.003

Fraile, J., Pardo, R., & Panadero, E. (2017). ¿Cómo emplear las rúbricas para implementar una verdadera evaluación formativa? *Revista Complutense de Educación*, *28*(4), 1321–1334. doi:10.5209/RCED.51915

Francis, J. E. (2018). Linking Rubrics and Academic Performance: An Engagement Theory Perspective. *Journal of University Teaching & Learning Practice*, *15*(1), 5–22. https://ro.uow.edu.au/jutlp/vol15/iss1/3. doi:10.53761/1.15.1.3

Frey, N., Fisher, D., & Hattie, J. (2018). Developing "assessment capable" learners. *Educational Leadership*, *75*(5), 46–51.

Frischer, J. and Larsson, K., (2000) Laissez-faire in research education—an inquiry into a Swedish doctoral program. *Higher Education Policy, 13*(2), 131-155.

Gabriel, S. (October, 2017). Moving from Silos and Burnout to Community and Engagement. *Faculty Focus.* Higher Ed Teaching Strategies from Magna Publications.

Galguera, T. (2011). Participant structures as professional learning tasks and the development of pedagogical language knowledge among preservice teachers. *Teacher Education Quarterly*, *38*(1), 85–106.

Gall, M. D., Gall, J. P., & Borg, W. T. (2003). *Educational research* (7th ed.). Pearson Education.

Garavan, T. N., & McGuire, D. (2001). Competencies and workplace learning: Some reflections on the rhetoric and the reality. *Journal of Workplace Learning*, *13*(4), 144–164. doi:10.1108/13665620110391097

Garcez-Manzanera, A. (2022). The affordances of rubrics in L2 writing in higher education: A new approach to enhancing writing conventions. *HUMAN Review*, *11*(Monográfico), 2–12. doi:10.37467/revhuman.v11.4086

García, O., & Wei, L. (2015). Translanguaging, bilingualism, and bilingual education. The handbook of bilingual and multilingual education, 223-240.

Garrett, N., Marques, J., & Dhiman, S. (2012). Assessment of business programs: A review of two models. *Business Education & Accreditation*, *4*(2), 17–25.

Gass, S. M., & Mackey, A. (2016). *Simulated recall methodology in applied linguistics and L2 research* (2nd ed.). Routledge. doi:10.4324/9781315813349

Gatfield, T. (2005). An investigation into PhD supervisory management styles: Development of a dynamic conceptual model and its managerial implications. *Journal of Higher Education Policy and Management*, *27*(3), 311–325. doi:10.1080/13600800500283585

Gebhard, M. (2010). Teacher education in changing times: A systemic functional linguistics (SFL) perspective. *TESOL Quarterly*, *44*(4), 797–803. doi:10.5054/tq.2010.237335

Gebhard, M., Chen, I.-A., & Britton, L. (2014). "Miss, nominalization is a nominalization:" English language learners' use of SFL metalanguage and their literacy practices. *Linguistics and Education*, *26*, 106–125. doi:10.1016/j.linged.2014.01.003

Gebhard, M., Demers, J., & Castillo-Rosenthal, Z. (2008). Teachers as critical text analysts: L2 literacies and teachers' work in the context of high-stakes school reform. *Journal of Second Language Writing*, *17*(4), 274–291. doi:10.1016/j.jslw.2008.05.001

General Medical Council. (2023). *Good medical practice*. GMC. www.gmc-uk.org/guidance

Giacumo, L. A., & Savenye, W. (2020). Asynchronous discussion forum design to support cognition: Effects of rubrics and instructor prompts on learner's critical thinking, achievement, and satisfaction. *Educational Technology Research and Development*, *68*(1), 37–66. doi:10.100711423-019-09664-5

Gibb, A. (2002). In pursuit of a new 'enterprise' and 'entrepreneurship' paradigm for learning: Creative destruction, new values, new ways of doing things and new combinations of knowledge. *International Journal of Management Reviews*, *4*(3), 233–269. doi:10.1111/1468-2370.00086

Gibbons, P. (2003). Mediating language learning: Teacher interactions with ESL students in a content-based classroom. *TESOL Quarterly*, *37*(2), 247–273. doi:10.2307/3588504

Gibbs, G. (1988). Learning by Doing : A Guide to Teaching and Learning Methods. *Further Education Unit*. [REMOVED HYPERLINK FIELD]

Gibbs, G., & Dunbar-Goddet, H. (2009). Characterising programme-level assessment environments that support learning. *Assessment & Evaluation in Higher Education*, *34*(4), 481–489. doi:10.1080/02602930802071114

Gibbs, G., Rust, C., Jenkins, A., & Jacques, D. (1994). *Developing students' transferable skills*. The Oxford Centre for Staff Development.

Gikandi, J. W., & Morrow, D. (2016). Designing and implementing peer formative feedback within online learning environments. *Technology, Pedagogy and Education*, *25*(2), 153–170. doi:10.1080/1475939X.2015.1058853

Gipps, C. (1999). Socio-cultural aspects of assessment. In A. Iran-Nejad & P. D. Pearson (Eds.), *Toward a new science of educational practice. Review of research in education* (Vol. 24, pp. 355–392). American Educational Research Association.

Giroux, H. A. (1988). *Teachers as intellectuals: Toward a critical pedagogy of learning*. Greenwood Publishing Group.

Glass, K. T. (2004). *Curriculum design for writing instruction: Creating standards-based lesson plans and rubrics*. Corwin Press.

Gogus, A. (2012). *Bloom's taxonomy of learning objectives. Encyclopedia of the Sciences of Learning*. Springer., doi:10.1007/978-1-4419-1428-6_141

González-Calvo, G., Varea, V., & Martínez-Álvarez, L. (2020). 'I feel, therefore I am': Unpacking preservice physical education teachers' emotions. *Sport Education and Society*, *25*(5), 543–555. doi:10.1080/13573322.2019.1620202

Gordon, S., & Smith, E. (2021). Who are faculty assessment leaders? *Assessment & Evaluation in Higher Education*, 1–14.

Graham, A., Harner, C., & Marsham, S. (2022). Can assessment specific marking criteria and electronic comment libraries increase student engagement with assessment and feedback? *Assessment & Evaluation in Higher Education*, *47*(7), 1071–1086. doi:10.1080/02602938.2021.1986468

Grainger, P. (2021). Enhancing assessment literacies through development of quality rubrics using a Triad based peer review process. *Journal of University Teaching & Learning Practice*, *18*(4), 4–14. doi:10.53761/1.18.4.4

Grainger, P., Purnell, K., & Zipf, R. (2008). Judging quality through substantive conversations between markers. *Assessment & Evaluation in Higher Education*, *33*(2), 133–142. doi:10.1080/02602930601125681

Grainger, P., & Weir, K. (2020). Creating quality rubrics through conversation. In P. Grainger & K. Weir (Eds.), *Facilitating student learning and engagement in higher education through assessment rubrics*. Cambridge Scholars Publishing.

Greenfield, B., Bridges, P., Phillips, T., Adams, E., Bullock, D., Davis, K., Nelson, C., & Wood, B. (2015). Reflective Narratives by Physical Therapist Students on Their Early Clinical Experiences: A Deductive and Inductive Approach.[]. *Journal, Physical Therapy Education*, *29*(2), 21–31. https://journals.lww.com/jopte. doi:10.1097/00001416-201529020-00005

Grierson, L., Winemaker, S., Taniguchi, A., Howard, M., Marshall, D., & Zazulak, J. (2020). The reliability characteristics of the REFLECT rubric for assessing reflective capacity through expressive writing assignments: A replication study. *Perspectives on Medical Education*, *9*(5), 281–285. doi:10.1007/S40037-020-00611-2 PMID:32803530

Gültekin, C. (2009). *Examining 9th grade students' abilities on drawing reading and interpreting of graphs about solutions and their properties.* [Unpublished master's thesis, Balikesir University, Türkiye].

Güneş, P., & Kılıç, D. (2016). Self-, peer- and teacher-assessment through rubrics. *Mehmet Akif Ersoy University Journal of Education Faculty*, *39*, 58–69.

Güneş, P., & Soran, H. (2013). Secondary school students' opinions on rubrics. *Kastamonu Education Journal*, *21*(4), 1327–1344.

Gurr, G. M. (2001). Negotiating the" Rackety Bridge"—A dynamic model for aligning supervisory style with research student development. *Higher Education Research & Development*, *20*(1), 81–92. doi:10.1080/07924360120043882

Gustafsson-Pesonen, A., & Remes, L. (2012). Evaluation of entrepreneurial development coaching: Changing the Teachers' thinking and action on entrepreneurship. *Annals of Innovation amp; Entrepreneurship*, *3*(1), 17211. doi:10.3402/aie.v3i0.17292

Hafner, J. C., & Hafner, P. M. (2003). Quantitative analysis of the rubric as an assessment tool: An emprical study of student peer-group rating. *International Journal of Science Education*, *25*(12), 1509–1528. doi:10.1080/0950069022000038268

Haladyna, T. M. (1997). *Writing test items to evaluate higher order thinking*. Allyn & Bacon.

Hale, J. A. (2008). *A guide to curriculum mapping*. Corwin Press.

Hall, C. M. (1991). Tourism as the Subject of Post-graduate Dissertations in Australia. *Annals of Tourism Research*, *18*(3), 520–523. doi:10.1016/0160-7383(91)90061-F

Hall, E. K., & Salmon, S. J. (2003). Chocolate chip cookies and rubrics helping students understand rubrics in inclusive settings. *Teaching Exceptional Children*, *35*(4), 8–11. doi:10.1177/004005990303500401

Halliday, M. A. (2013). Meaning as choice. In L. Fontaine, Tom Bartlett, & G. O'Grady (Eds.), Systemic functional linguistics: Exploring choice (pp. 15-36). Cambridge University Press. doi:10.1017/CBO9781139583077.003

Halliday, M. A. K. (1993). Towards a language-based theory of learning. *Linguistics and Education*, *5*(2), 93–116. doi:10.1016/0898-5898(93)90026-7

Halliday, M. A. K., & Hasan, R. (1989). *Language, context, and text: Aspects of language in a social-semiotic perspective*. Oxford University Press.

Halliday, M. A. K., & Kress, G. R. (1976). *System and function in language: Selected papers*. Oxford University Press.

Halliday, M. A. K., & Matthiessen, C. (2014). *An introduction to functional grammar*. Edward Arnold. doi:10.4324/9780203783771

Halquist, D., & Musanti, S. I. (2010). Critical incidents and reflection: Turning points that challenge the researcher and create opportunities for knowing. *International Journal of Qualitative Studies in Education : QSE*, *23*(4), 449–461. doi:10.1080/09518398.2010.492811

Hämäläinen, M., Ruskovaara, E., & Pihkala, T. (2018). Principals promoting entrepreneurship education: The relationships between development activities and school practises. *Journal of Entrepreneurship Education*, *21*(2), 1–19.

Hamel, J., Dufour, S., & Fortin, D. (1993). *Case Study Methods*. SAGE. doi:10.4135/9781412983587

Hamodi, C., & Barba-Martín, R. A. (2021). *Evaluación formativa y compartida: nuevas propuestas de desarrollo en Educación Superior*. Dextra.

Hamodi, C., Moreno, J., & Barba-Martín, R. A. (2018). Medios de evaluación y desarrollo de competencias en educación superior en estudiantes de educación física. *Estudios Pedagógicos (Valdivia)*, *44*(2), 241–257. doi:10.4067/S0718-07052018000200241

Handley, K., Price, M., & Millar, J. (2011). Beyond "doing time": Investigating the concept of student engagement with feedback. *Oxford Review of Education*, *37*(4), 543–560. https://www.jstor.org/stable/23047914. doi:10.1080/03054985.2011.604951

Harden, R. M., Sowden, S., & Dunn, W. R. (1984). Educational strategies in curriculum development: The SPICES model. *Medical Education*, *18*(4), 284–297. doi:10.1111/j.1365-2923.1984.tb01024.x PMID:6738402

Harland, T., & Wald, N. (2021). The assessment arms race and the evolution of a university's assessment practices. *Assessment & Evaluation in Higher Education*, *46*(1), 105–117. doi:10.1080/02602938.2020.1745753

Hattie, J., & Timperley, H. (2007). The power of feedback. *Review of Educational Research*, *77*(1), 81–112. doi:10.3102/003465430298487

Hattikudur, S., Prather, R. W., Asquith, P. S., Alibali, M. W., Knuth, E. J., & Nathan, M. (2012). Constructing graphical representations: Middle schoolers' intuitions and developing knowledge about slope and y-intercept. *School Science and Mathematics*, *112*(3), 230–240. doi:10.1111/j.1949-8594.2012.00138.x

Hatton, N., & Smith, D. (1995). Reflection in teacher education: Towards definition and implementation. *Teaching and Teacher Education*, *11*(1), 33–49. doi:10.1016/0742-051X(94)00012-U

Hauk, S., Toney, A., Jackson, B., Nair, R., & Tsay, J. J. (2014). Developing a Model of Pedagogical Content Knowledge for Secondary and Post-Secondary Mathematics Instruction. *Dialogic Pedagogy, 2*.

Hawe, E., Dixon, H., Murray, J., & Chandler, S. (2021). Using rubrics and exemplars to develop students' evaluative and productive knowledge and skill. *Journal of Further and Higher Education*, *45*(8), 1033–1047. doi:10.1080/03098 77X.2020.1851358

Haworth, R. (2016). Personal Learning Environments: A solution for self-directed learners. *TechTrends*, *60*(4), 359–364. doi:10.100711528-016-0074-z

Hawthorne, K. A., Bol, L., & Pribesh, S. (2017). Can providing rubrics for writing tasks improve developing writers' calibration accuracy? *Journal of Experimental Education*, *85*(4), 689–708. doi:10.1080/00220973.2017.1299081

Hayes, A. F., & Krippendorff, K. (2007). Answering the call for a standard reliability measure for coding data. *Communication Methods and Measures*, *1*(1), 77–89. doi:10.1080/19312450709336664

Hay, P., & Penney, D. (2013). *Assessment in Physical Education: A Socio-cultural Higher Education*. Australian Learning and Teaching Council., Available at http://www.uts.edu.au/sites/default/files/ Assessment-2020_propositions_final.pdf

Healey, M., Flint, A., & Harrington, K. (2014). *Engagement through partnership: students as partners in learning and teaching in Higher Education*. Academy HE. https://www.heacademy.ac.uk/engagement-through-partnership-students-partners-learning-and-teaching-higher-education

Healey, M., Flint, A., & Harrington, K. (2016). Students As Partners: Reflections on a Conceptual Model. *Teaching & Learning Inquiry*, *4*(2), 8–20. doi:10.20343/teachlearninqu.4.2.3

Healey, M., & Healey, R. L. (2019). *Student engagement through partnership: A guide and update to the Advance HE framework*. Advance HE. mickhealey.co.uk/devtest/wp-content/ uploads/2020/02/SaP-Guide-Updated.pdf

Health and Care Professions Council. (2023). *Standards*. HCPC. https://www.hcpc-uk.org/standards/

Henry, C., Hill, F., & Leitch, C. (2005). Entrepreneurship education and training: Can entrepreneurship be taught? Part I. *Education + Training*, *47*(2), 98–111. doi:10.1108/00400910510586524

Henry, C., & Lewis, K. (2018). A review of entrepreneurship education research: Exploring the contribution of the Education + Training special issues. *Education + Training*, *60*(3), 263–286. doi:10.1108/ET-12-2017-0189

Heras Bernardino, C., & Herrán Álvarez, I. (2019). La evaluación formativa y compartida desde un enfoque competencial. Aplicación práctica en tareas de aula en Primaria y Secundaria. *Revista Infancia, Educación Y Aprendizaje, 5*(2), 568–575. doi:10.22370/ieya.2019.5.2.1777

Hernandez, J., & Korzun, B. (2021). Methods to select the scales and starting values for axes in linear graphs. *The Physics Teacher*, *59*(6), 482–483. doi:10.1119/10.0006136

Hernandez, K. A. C., Chang, H., & Ngunjiri, F. W. (2017). Collaborative autoethnography as multivocal, relational, and democratic research: Opportunities, challenges, and aspirations. *a/b*. *Auto/Biography Studies: a/B*, *32*(2), 251–254. doi:10.1080/08989575.2017.1288892

Hernando-Garijo, A., Hortigüela-Alcalá, D., & Pérez-Pueyo, Á. (2017). El proceso de evaluación formativa en la realización de un vídeo tutorial de estiramientos en inglés en un centro bilingüe. In V. M. López-Pastor & Á. Pérez-Pueyo (Eds.), *Evaluación Formativa y Compartida en Educación: experiencias de éxito en todas las etapas educativas* (pp. 260–270). Universidad de León.

Hernán, E., López-Pastor, V., & Pérez-Brunicardi, D. (2019). Por qué hago evaluación formativa y compartida y/o evaluación para el aprendizaje en EF? La influencia de la formación inicial y permanente del profesorado. *Revista RETOS*, *36*, 37–43.

Hill, J., Berlin, K., Choate, J., Cravens-Brown, L., McKendrick-Calder, L., & Smith, S. (2021). Can Relational Feed-Forward Enhance Students' Cognitive and Affective Responses to Assessment? *Teaching & Learning Inquiry*, *9*(2), n2. doi:10.20343/teachlearninqu.9.2.18

Hodgson, K., Lamport, D., & Laville, A. (2021). Variable trajectory: a systematic review, analytic synthesis and construct domain consolidation of international measures of competence in doctors and medical students. *BMJ Open, 11* (8). doi:10.1136/bmjopen-2020-047395

Holmes, C., & Oakleaf, M. (2013). The official (and unofficial) rules for norming rubrics successfully. *Journal of Academic Librarianship*, *39*(6), 599–602. doi:10.1016/j.acalib.2013.09.001

Honeychurch, S. (2015). *Rubrics as a guide to student writing and staff grading*. Chartered Association of Business Schools.

Hortigüela-Alcalá, D., Abella García, V., & Pérez-Pueyo, Á. (2015). ¿ De qué manera se implica el alumnado en el aprendizaje? Análisis de su percepción en procesos de evaluación formativa. *Revista de Investigación Educacional*, *13*(1), 88–104.

Hortigüela-Alcalá, D., González-Víllora, S., & Hernando-Garijo, A. (2021). Do we really assess learning in Physical Education? Teacher's perceptions at different educational stages. *Retos*, *42*, 643–654. doi:10.47197/retos.v42i0.88686

Hortigüela-Alcalá, D., Palacios, A., & López-Pastor, V. (2019a). The impact of formative and shared or coassessment on the acquisition of transversal competences in higher education. *Assessment & Evaluation in Higher Education*, *44*(6), 933–945. doi:10.1080/02602938.2018.1530341

Hortigüela-Alcalá, D., Pérez-Pueyo, A., & González-Calvo, G. (2019b). Pero… ¿A qué nos referimos realmente con la Evaluación Formativa y Compartida?: Confusiones Habituales y Reflexiones Prácticas. *Revista Iberoamericana de Evaluación Educativa*, *12*(1), 13–27. doi:10.15366/riee2019.12.1.001

Huang, S.-C. (2012). Like a bell responding to a striker: Instruction contingent on assessment. *English Teaching*, *11*(4), 99–119.

Huba, M. E., & Freed, J. E. (2000). Using rubrics to provide feedback to students. Learner-Centered Assessment on College Campuses. Allyn and Bacon, 151–200.

Huball, H., & Burt, H. (2004). An Integrated Approach to Developing and Implementing Learning Centred Curricula. *The International Journal for Academic Development*, *9*(1), 51–65. doi:10.1080/1360144042000296053

Hughes, A. L., Michener, C., Mohamed, K., & Mcduff, N. (2019). Curriculum co-creation as a transformative strategy to address differential student outcomes: The example of Kingston University's Student Curriculum Consultant Programme. *Compass (Eltham)*, *12*(1), 1. Advance online publication. doi:10.21100/compass.v12i1.955

Hull, M. M., Kuo, E., Gupta, A., & Elby, A. (2013). Problem-solving rubrics revisited: Attending to the blending of informal conceptual and formal mathematical reasoning. *Physical Review Special Topics- Review Physics*. *Education Research*, *9*, 010105.

Hyland, F. (1998). The impact of teacher written feedback on individual writers. *Journal of Second Language Writing*, *7*(3), 255–286. doi:10.1016/S1060-3743(98)90017-0

Hyland, K. (1996). Writing without conviction? Hedging in science research articles. *Applied Linguistics*, *17*(4), 433–454. doi:10.1093/applin/17.4.433

İnanç, H. (2019). Graphic Drawing skills of science teaching candidates. *International Scientific and Vocational Journal*, *3*(1), 22–30.

Inoue, A. B. (2015). *Antiracist writing assessment ecologies: Teaching and assessing writing for a socially just future*. Parlor Press LLC. doi:10.37514/PER-B.2015.0698

International Association for Physical Education in Higher Education (2021). Tomada de Posição Sobre Avaliação em Educação Física. *Boletim SPEF, 42*.

Jacobs, H. (1997). *Mapping the big picture: Integrating curriculum and assessment K-12*. Association for Supervision and Curriculum Development.

Jeong, H. (2015). Rubrics in the classroom: Do teachers really follow them? *Language Testing in Asia*, *5*(1), 1–14. doi:10.118640468-015-0013-5

Jerusalem, M., & Schwarzer, R. (1995). *General Self-Efficacy Scale—Revised—English Version (Gse-R)* [Database record]. APA PsycTests. doi:10.1037/t18916-000

Jessop, T., Hakim, Y., & Gibbs, G. (2014). The whole is greater than the sum of its parts: A large-scale study of students' learning in response to different programme assessment patterns. *Assessment & Evaluation in Higher Education*, *39*(1), 39. doi:10.1080/02602938.2013.792108

Jessop, T., & Tomas, C. (2016). The implications of programme assessment patterns for student learning. *Assessment & Evaluation in Higher Education*.

Johansen, V., & Schanke, T. (2013). Entrepreneurship education in secondary education and training. *Scandinavian Journal of Educational Research*, *57*(4), 357–368. doi:10.1080/00313831.2012.656280

Johns, C. & Freshwater, D. (1998). *Transforming nursing through reflective practice*. 222.

Johnston, P. H. (2004). *Choice words: How our language affects children's learning*. Stenhouse Publishers.

Johnston, P., & Andrade, H. (2012). Assessment, teaching and learning in and beyond classrooms. In B. Kaur (Ed.), *Understanding teaching and learning: Classroom research revisited* (pp. 269–280). Sense Publishers. doi:10.1007/978-94-6091-864-3_20

Jones, B., & Iredale, N. (2010). Enterprise education as pedagogy. *Education + Training*, *1*(52), 7–19. doi:10.1108/00400911011017654

Jones, L., Allen, B., Dunn, P., & Brooker, L. (2017). Demystifying the rubric: A five-step pedagogy to improve student understanding and utilisation of marking criteria. *Higher Education Research & Development*, *36*(1), 129–142. doi:10.1080/07294360.2016.1177000

Jonsson, A. (2014). Rubrics as a way of providing transparency in assessment. *Assessment & Evaluation in Higher Education*, *39*(7), 840–852. doi:10.1080/02602938.2013.875117

Jönsson, A., & Panadero, E. (2017). The Use and Design of Rubrics to Support Assessment for Learning. In D. Carless, S. M. Bridges, C. K. Y. Chan, & R. Glofcheski (Eds.), *Scaling up Assessment for Learning in Higher Education* (pp. 99–111). Springer. doi:10.1007/978-981-10-3045-1_7

Jonsson, A., & Svingby, G. (2007). The use of scoring rubrics: Reliability, validity and educational consequences. *Educational Research Review*, *2*(2), 130–144. doi:10.1016/j.edurev.2007.05.002

Joseph, S., Rickett, C., Northcote, M., & Christian, B. J. (2020). 'Who are you to judge my writing?': Student collaboration in the co-construction of assessment rubrics. *New Writing*, *17*(1), 31–49. doi:10.1080/14790726.2019.1566368

Joyner, H. S. (2016). Curriculum mapping: A method to assess and refine undergraduate degree programs. *Journal of Food Science Education*, *15*(3), 83–100. doi:10.1111/1541-4329.12086

Jutras, D. (2023). *Scales, stars and numbers: the question of evaluation*. Times Higher Education.

Kahu, E. R., & Nelson, K. (2018). Student engagement in the educational interface: Understanding the mechanisms of student success. *Higher Education Research & Development*, *37*(1), 58–71. doi:10.1080/07294360.2017.1344197

Kan, A. (2009). Ödev ve Projeler. Eğitimde Ölçme ve Değerlendirme, Hakan Atılgan (Ed.), 4. Baskı Anı Yayıncılık, Ankara.

Karaçam, Z. (2013). Systematic review methodology: A guide for preparation of systematic review. *E-Journal of Dokuz Eylul University Nursing Faculty*, *6*(1), 26–33.

Karal Eyüboğlu, I. S. (2020). Interpretation of an energy graph for a mass-spring system by prospective science and mathematics teachers: A comparison. *Online Science Education Journal*, *5*(2), 52–59.

Karslı, M. B., Karabey, S. Ç., Nergiz, E., & Goktas, Y. (2018). Comparison of the discussion sections of PhD dissertations in educational technology: The case of Turkey and the USA. *Scientometrics*.

Kearney, S. (2013). Improving Engagement: The Use of 'Authentic Self-and Peer-assessment for Learning' to Enhance the Student Learning Experience. *Assessment & Evaluation in Higher Education*, *38*(7), 875–891. doi:10.1080/02602938.2012.751963

Kember, D. A. (1997). A reconceptualisation of the research into university academics' conceptions of teaching. *Learning and Instruction*, *7*(3), 255–275. doi:10.1016/S0959-4752(96)00028-X

Khanmohammad, H., & Osanloo, M. (2009). *Moving toward objective scoring: A rubric for translation assessment*.

Kilgour, A., Morton, J., Cloete, L., Dawson, S., & Northcote, M. (2022). *Rubric co-construction in medical and allied health education: Students' and teachers' perceptions* [Preprint]. In Review. doi:10.21203/rs.3.rs-1984776/v1

Kilgour, P., Northcote, M., Williams, A., & Kilgour, A. (2020). A Plan for the Co-Construction and Collaborative Use of Rubrics for Student Learning. *Assessment & Evaluation in Higher Education*, *45*(1), 140–153. doi:10.1080/02602938.2019.1614523

Kim, K. H., & Zabelina, D. (2015). Cultural bias in assessment: Can creativity assessment help? *The International Journal of Critical Pedagogy*, *6*(2).

Kirk, R. E. (2007). *Statistics, an introduction* (5th ed.). Thomson Wadsworth.

Knoch, U. (2009). Diagnostic assessment of writing: A comparison of two rating scales. *Language Testing*, *26*(2), 275–304. doi:10.1177/0265532208101008

Kocakülah, A. (2022). Development and use of a rubric to assess undergraduates' problem solutions in Physics. *Participatory Educational Research*, *9*(3), 362–382. doi:10.17275/per.22.71.9.3

Kocakülah, M. S. (2010). Development and application of a rubric for evaluating students' performance on Newton's laws of motion. *Journal of Science Education and Technology*, *19*(2), 46–164. doi:10.100710956-009-9188-9

Kocakülah, M. S., & Aytaç, N. N. (2007). Development and use of a rubric for the assessment of students' performance in solving problems on Newton's law of motion. *AIP Conference Proceedings*, *899*, 838. doi:10.1063/1.2733579

Koh, K. H. (2011). Improving teachers' assessment literacy through professional development. *Teaching Education*, *22*(3), 255–276. doi:10.1080/10476210.2011.593164

Kohn, A. (2006). The trouble with rubrics. *English Journal*, *95*(4), 12–15. doi:10.2307/30047080

Kolb, D. A. (1984). *Experiential learning : experience as the source of learning and development*. Elsevier.

Komarkova, I., Conrads, J., & Collado, A. (2015). Entrepreneurship Competence: An Overview of Existing Concepts. *Policies and Initiatives. depth case study report*. https://citeseerx.ist.psu.edu/viewdoc/download?doi=10.1.1.874.7140 &rep=rep1&type=pdf

Krebs, R., Rothstein, B., & Roelle, J. (2022). Rubrics enhance accuracy and reduce cognitive load in self-assessment. *Metacognition and Learning*, *17*(2), 627–650. doi:10.100711409-022-09302-1

Kremer, L., & Hofman, J. E. (1979). A Three-Dimensional Typology of Teacher Personality. *The Journal of Educational Research*, *73*(1), 20–25. doi:10.1080/00220671.1979.10885198

Krippendorff, K. (2004). Reliability in content analysis: Some common misconceptions and recommendations. *Human Communication Research*, *30*(3), 411–433. doi:10.1111/j.1468-2958.2004.tb00738.x

Kubiszyn, T., & Borich, G. (2013). *Educational testing and measurement*. John Wiley & Sons, Inc.

Kujala, S. (2003). User involvement: A review of the benefits and challenges. *Behaviour & Information Technology*, *22*(1), 1–16. doi:10.1080/01449290301782

Kuntze, S. (2012). Pedagogical content beliefs: Global, content domain-related and situation-specific components. *Educational Studies in Mathematics*, *79*(2), 273–292. doi:10.100710649-011-9347-9

Kwon, O. N. (2002). The effect of calculator based ranger activities on students' graphing ability. *School Science and Mathematics*, *102*(2), 57–67. doi:10.1111/j.1949-8594.2002.tb17895.x

Kyndt, E., & Baert, H. (2015). Entrepreneurial competencies: Assessment and predictive value for entrepreneurship. *Journal of Vocational Behavior*, *90*, 13–25. doi:10.1016/j.jvb.2015.07.002

Kyvik, S., & Thune, T. (2015). Assessing the quality of PhD dissertations. A survey of external committee members. *Assessment & Evaluation in Higher Education*, *40*(5), 768–782. doi:10.1080/02602938.2014.956283

Lackéus, M., & Sävetun, C. (2019). Assessing the impact of enterprise education in three leading Swedish compulsory schools. *Journal of Small Business Management*, *57*(sup1), 33–59. doi:10.1111/jsbm.12497

Lafuente, J. C., & Hortigüela-Alcalá, D. (2021). La percepción de los futuros maestros respecto a la implantación de contenidos de expresión corporal. *Movimento (Porto Alegre)*, *27*, 1–15. doi:10.22456/1982-8918.111735

Landis, J. R., & Koch, G. G. (1977). The measurement of observer agreement for categorical data. *Biometrics*, *33*(1), 159–174. doi:10.2307/2529310 PMID:843571

Lans, T., & Gulikers, J. (2010). Assessing entrepreneurial competence in entrepreneurship education and training. Handbook of research in entrepreneurship education, 3, 54-67.

Laville, A., Chessell, C., & Wiehe, T. (n.d.) Developing and embedding electronic assessment overviews. Teaching and Learning Exchange, CQSD: University of Reading.

Laville, A., Holtom, H., Conway, E., & Alder, C. (2023). Exploring clinical perfectionism in higher education students: key recommendations and reflections on a partnership. *International Journal for Students as Partners, 7* (1).

Lawshe, C. H. (1975). A quantitative approach to content validity. *Personnel Psychology*, *28*(4), 563–575. doi:10.1111/j.1744-6570.1975.tb01393.x

Le Deist-Delamare, F. D., & Winterton, J. (2005). What is competence? *Human Resource Development International*, *8*(1), 27–46. doi:10.1080/1367886042000338227

Leader, D. C., & Clinton, M. S. (2018). Student perceptions of the effectiveness of rubrics. *Journal of Business and Educational Leadership*, *8*(1), 86–103.

Lea, M. R., & Street, B. V. (1998). Student writing in higher education: An academic literacies approach. *Studies in Higher Education*, *23*(2), 157–172. doi:10.1080/03075079812331380364

Lee, A. (2012). *Successful research supervision: Advising students doing research.* Routledge. doi:10.4324/9780203816844

Lee, J. (2023). *Effective assessment practices for a ChatGPT-enabled world.* Times Higher Education.

Leenknecht, M., Wijnia, L., Köhlen, M., Fryer, L., Rikers, R., & Loyens, S. (2021). Formative assessment as practice: The role of students' motivation. *Assessment & Evaluation in Higher Education*, *46*(2), 236–255. doi:10.1080/02602938.2020.1765228

Lee, Y. W., Gentile, C., & Kantor, R. (2010). Toward automated multi-trait scoring of essays: Investigating links among holistic, analytic, and text feature scores. *Applied Linguistics*, *31*(3), 391–417. doi:10.1093/applin/amp040

Leisen, M. (2022). Make your rubric more than a wall of words. *Educational Leadership*, *79*(7). https://www.ascd.org/el/articles/make-your-rubric-more-than-a-wall-of-words

Leist, C. W., Woolwine, M. A., & Bays, C. L. (2012). The effects of using a critical thinking scoring rubric to assess undergraduate students' reading skills. *Journal of College Reading and Learning*, *43*(1), 31–58. doi:10.1080/10790195.2012.10850361

Leki, I. (1995). Coping strategies of ESL students in writing tasks across the curriculum. *TESOL Quarterly*, *29*(2), 235–260. doi:10.2307/3587624

Leki, I. (2001). "A narrow thinking system": nonnative-English-speaking students in group projects across the curriculum. *TESOL Quarterly*, *35*(1), 39–67. doi:10.2307/3587859

Lewis, E. (2021). Best Practices for Improving the Quality of the Online Course Design and Learners Experience. *The Journal of Continuing Higher Education*, *69*(1), 61–70. doi:10.1080/07377363.2020.1776558

Lewis, J. (1877). The rubric: its history and meaning. In *The Oxford Movement: Tractarian Pamphlets at Pusey House: The Halifax and Church Sub-Collections* (2nd ed.). Adams, and Co.

Libarkin, J. (2008). *Concept inventories in higher education science.* Manuscript prepared for the National Research Council Promising Practices in Undergraduate STEM Education Workshop 2, Washington, D.C.

Lieberman, M. D. (2012). What zombies can't do: A social cognitive neuroscience approach to the irreducibility of reflective consciousness. In *Two Minds*. Dual Processes and Beyond. doi:10.1093/acprof:oso/9780199230167.003.0013

Liew, S. S., Lim, H. L., Saleh, S., & Ong, S. L. (2019). Development of scoring rubrics to assess physics practical skills. *Eurasia Journal of Mathematics, Science and Technology Education*, *15*(4), 1–14. doi:10.29333/ejmste/103074

Li, J., & Lindsey, P. (2015). Understanding variations between student and teacher application of rubrics. *Assessing Writing*, *26*, 67–79. doi:10.1016/j.asw.2015.07.003

Lilleväli, U., & Täks, M. (2017). *Competence models as a tool for conceptualizing the systematic process of entrepreneurship competence development.* Education Research International. doi:10.1155/2017/5160863

Lim, J. Y., Yew, S., Ong, K., Yan, C., Ng, H., Li, K., Chan, E., Yi, S., Wu, E. A., So, W. Z., Jin, G., Tey, C., Lam, Y. X., Lu, N., Gao, X., Lim, Y. X., Yong, R., Tay, K., Tze, I., & Krishna, R. (2023). A systematic scoping review of reflective writing in medical education. *BMC Medical Education*, *23*(12), 12. doi:10.118612909-022-03924-4 PMID:36624494

Lin, Y. C. (2012). *Online supervision of school counselors: Effects on case conceptualization skills and self-efficacy* [Doctoral dissertation, The University of Iowa].

Lindström, L. (2006). Creativity: What is it? Can you assess it? Can it be taught? *International Journal of Art & Design Education*, *25*(1), 53–66. doi:10.1111/j.1476-8070.2006.00468.x

Linnenbrink, E. A., & Pintrich, P. R. (2003). The Role of Self-Efficacy Beliefs In Student Engagement and Learning In the classroom. *Reading & Writing Quarterly*, *19*(2), 119–137. doi:10.1080/10573560308223

Lipnevich, A. A., McCallen, L. N., Miles, K. P., & Smith, J. K. (2014). Mind the gap! Students' use of exemplars and detailed rubrics as formative assessment. *Instructional Science*, *42*(4), 539–559. doi:10.100711251-013-9299-9

Logan, D., King, J., & Fischer-Wright, H. (2011). *Tribal Leadership: Leveraging Natural Groups to Build a Thriving Organization*. Harper Business.

López-Pastor, V. M., Kirk, D., Lorente-Catalán, E., MacPhail, A., & Macdonald, D. (2013). Alternative assessment in physical education: A review of international literature. *Sport Education and Society*, *18*(1), 57–76. doi:10.1080/1357 3322.2012.713860

López-Pastor, V. M., & Pérez-Pueyo, Á. (2017). *Evaluación formativa y compartida en educación: experiencias de éxito en todas las etapas educativas*. Universidad de León.

Lorber, P., Rooney, S., & Van Der Enden, M. (2019). Making assessment accessible: A student–staff partnership perspective. *Higher Education Pedagogies*, *4*(1), 488–502. doi:10.1080/23752696.2019.1695524

Lorente-Catalán, E., & Kirk, D. (2015). Student teachers' understanding and application of assessment for learning during physical education teacher education course. *European Physical Education Review*, *22*(1), 65–81. doi:10.1177/1356336X15590352

Lovitts, B. E. (2005). How to grade a dissertation. *Academe, 91*(6), 18-23.

Lowrie, T., & Diezmann, C. M. (2007). Middle school students interpreting graphical tasks: difficulties within a graphical language. In *4th East Asia Regional Conference on Mathematics Education*, (pp. 611-617).

Lucas, C., Smith, L., Lonie, J. M., Hough, M., Rogers, K., & Mantzourani, E. (2019). *Can a reflective rubric be applied consistently with raters globally? A study across three countries*. Science Direct. doi:10.1016/j.cptl.2019.06.004

Luft, J. A. (1999). Rubrics: Design and use in science teacher education. *Journal of Science Teacher Education*, *10*(2), 107–121. doi:10.1023/A:1009471931127

Lytle, S. L., & Cochran-Smith, M. (1994). Chapter II: Inquiry, Knowledge, and Practice. *Teachers College Record*, *95*(6), 22–51. doi:10.1177/016146819409500602

Mabry, L. (1999). Writing to the rubric: Lingering effects of traditional standardized testing on direct writing assessment. *Phi Delta Kappan*, *80*(9), 673–679.

Macken-Horarik, M. (2008). *A "good enough" grammatics: Developing an effective metalanguage for school English in an era of multiliteracies*. The ISFC, Sydney.

MacKenzie, W. (2004). Constructing a rubric. NETS'S curriculum series: Social studies units for grades 9–12, International Society for Technology in Education, 24–30.

MacLean, J. (2018). Teachers as Agents of Change in Curricular Reform: The Position of Dance Revisited. *Sport Education and Society*, *23*(6), 563–577. doi:10.1080/13573322.2016.1249464

Mainhard, T., Van Der Rijst, R., Van Tartwijk, J., & Wubbels, T. (2009). A model for the supervisor–doctoral student relationship. *Higher Education*, *58*(3), 359–373. doi:10.100710734-009-9199-8

Maloney, S., Tai, J. H. M., Lo, K., Molloy, E., & Ilic, D. (2013). Honesty in critically reflective essays: An analysis of student practice. *Advances in Health Sciences Education : Theory and Practice*, *18*(4), 617–626. doi:10.100710459-012-9399-3 PMID:22926807

Mama, M., & Hennessy, S. (2013). Developing a typology of teacher beliefs and practices concerning classroom use of ICT. *Computers & Education*, *68*, 380–387. doi:10.1016/j.compedu.2013.05.022

Mann, K., Gordon, J., & MacLeod, A. (2009). Reflection and reflective practice in health professions education: A systematic review. *Advances in Health Sciences Education : Theory and Practice*, *14*(4), 595–621. doi:10.100710459-007-9090-2 PMID:18034364

Manolescu, M. (2005). *Evaluarea scolara* [Lexical characteristics of Romanian language]. Meteor press.

Man, T. W. Y., Lau, T., & Chan, K. F. (2002). The competitiveness of small and medium enterprises: A conceptualisation with focus on entrepreneurial competencies. *Journal of Business Venturing*, *17*(2), 123–142. doi:10.1016/S0883-9026(00)00058-6

Markman, G. D., Baron, R. A., & Balkin, D. B. (2005). Are perseverance and self-efficacy costless? Assessing entrepreneurs' regretful thinking. *Journal of Organizational Behavior: The International Journal of Industrial*. *Journal of Organizational Behavior*, *26*(1), 1–19. doi:10.1002/job.305

Martens, K. S. (2018). How program evaluators use and learn to use rubrics to make evaluative reasoning explicit. *Evaluation and Program Planning*, *69*, 25–32. doi:10.1016/j.evalprogplan.2018.03.006 PMID:29660495

Martens, S. E., Spruijt, A., Wolfhagen, I. H. A. P., Whittingham, J. R. D., & Dolmans, D. H. J. M. (2019). A students' take on student-staff partnerships: Experiences and preferences. *Assessment & Evaluation in Higher Education*, *44*(6), 910–919. doi:10.1080/02602938.2018.1546374

Martin, J. R., & Rose, D. (2008). *Genre relations: Mapping culture*. Equinox.

Mastrorilli, T. M., Harnett, S., & Zhu, J. (2014). *Arts Achieve* impacting student success in the arts: Preliminary findings after one year of implementation. *Journal for Learning through the Arts, 10*(1). https://escholarship.org/uc/item/6c81239d

Matshedisho, K. (2020). Straddling rows and columns: Students' (mis)conceptions of an assessment rubric. *Assessment & Evaluation in Higher Education*, *45*(2), 169–179. doi:10.1080/02602938.2019.1616671

Mattsson, T., & Lundvall, S. (2015). The Position of Dance in Physical Education. *Sport Education and Society*, *20*(7), 855–871. doi:10.1080/13573322.2013.837044

May, C. R., Albers, B., Bracher, M., Finch, T. L., Gilbert, A., Girling, M., Greenwood, K., MacFarlane, A., Mair, F. S., May, C. M., Murray, E., Potthoff, S., & Rapley, T. (2022). Translational framework for implementation evaluation and research: A normalisation process theory coding manual for qualitative research and instrument development. *Implementation Science : IS*, *17*(1), 19. doi:10.118613012-022-01191-x PMID:35193611

May, C. R., Cummings, A., Girling, M., Bracher, M., Mair, F. S., May, C. M., Murray, E., Myall, M., Rapley, T., & Finch, T. (2018). Using Normalization Process Theory in feasibility studies and process evaluations of complex healthcare interventions: A systematic review. *Implementation Science : IS*, *13*(1), 80. doi:10.118613012-018-0758-1 PMID:29879986

Compilation of References

May, C. R., Mair, F., Finch, T., MacFarlane, A., Dowrick, C., Treweek, S., Rapley, T., Ballini, L., Ong, B. N., Rogers, A., Murray, E., Elwyn, G., Légaré, F., Gunn, J., & Montori, V. M. (2009). Development of a theory of implementation and integration: Normalization Process Theory. *Implementation Science : IS*, *4*(1), 29. doi:10.1186/1748-5908-4-29 PMID:19460163

McAlpine, M. (2002). *Principles of assessment*. University of Glasgow, Robert Clark Center for Technological Education., http://www.caacentre.ac.uk/dldocs/Bluepaper1.pdf

McDermott, L. C., Rosenquist, M. L., & Van Zee, E. H. (1987). Student difficulties in connecting graphs and physics: Examples from kinematics. *American Journal of Physics*, *55*(6), 503–513. doi:10.1119/1.15104

McDuff, N., Hughes, A., Tatam, J., Morrow, E., & Ross, F. (2020). Improving equality of opportunity in Higher Education through the adoption of an inclusive curriculum framework. *Widening Participation and Lifelong Learning : the Journal of the Institute for Access Studies and the European Access Network*, *22*(2), 83–121. doi:10.5456/WPLL.22.2.83

McGee, J. E., Peterson, M., Mueller, S. L., & Sequira, M. J. (2009). Entrepreneurial self-efficacy: Refining the measure. *Entrepreneurship Theory and Practice*, *33*(4), 965–988. doi:10.1111/j.1540-6520.2009.00304.x

Mckenzie, D. L., & Padilla, M. J. (1986). The construction and validation of the test of graphing in science (TOGS). *Journal of Research in Science Teaching*, *23*(7), 571–579. doi:10.1002/tea.3660230702

McKinney, B. (2018). The impact of program-wide discussion board grading rubrics on students and faculty satisfaction. *Online Learning : the Official Journal of the Online Learning Consortium*, *22*(2), 289–300. doi:10.24059/olj.v22i2.1386

McNair, T. B. (2016). Designing purposeful pathways for student achievement through transparency and problem-centered learning. *Peer Review : Emerging Trends and Key Debates in Undergraduate Education*, *18*(1–2), 3–6. https://go.gale.com/ps/i.do?p=AONE&sw=w&issn=15411389&v=2.1&it=r&id=GALE%7CA459505877&sid=googleScholar&linkaccess=abs

McTighe, J., & Frontier, T. (2022). How to provide better feedback through rubrics. *Educational Leadership*, *79*(7), 17–23.

Medland, E. (2019). 'I'm an assessment illiterate': Towards a shared discourse of assessment literacy for external examiners. *Assessment & Evaluation in Higher Education*, *44*(4), 565–580. doi:10.1080/02602938.2018.1523363

Menéndez-Varela, J. L., & Gregori-Giralt, E. (2016). The contribution of rubrics to the validity of performance assessment: A study of the conservation–restoration and design undergraduate degrees. *Assessment & Evaluation in Higher Education*, *41*(2), 228–244. doi:10.1080/02602938.2014.998169

Mercer-Mapstone, L., & Marie, J. (2019). *Practical guide: Scaling up student-staff partnerships in higher education*. Institute for Academic Development: University of Edinburgh. http://bit.ly/2EfUR16

Mercer-Mapstone, L., & Bovill, C. (2020). Equity and diversity in institutional approaches to student–staff partnership schemes in higher education. *Studies in Higher Education*, *45*(12), 2541–2557. doi:10.1080/03075079.2019.1620721

Mertler, C. A. (2001). Designing scoring rubrics for your classroom. *Practical Assessment, Research & Evaluation*, *7*(25), 1–8.

Mettler, C. A. (2002). Designing scoring rubrics for your classroom. *Understanding Scoring Rubrics: A Guide for Teachers, ERIC Clearinghouse on Assessment and Evaluation*, 72–81. ERIC.

Mijušković, M. (2014). Assessing students' reading comprehension through rubrics. *Mediterranean Journal of Social Sciences*, *5*(13), 252.

Miller-Kuhlmann, R., Osullivan, P. S., & Aronson, L. (2015). Essential steps in developing best practices to assess reflective skill: A comparison of two rubrics. Taylor & Francis. doi:10.3109/0142159X.2015.1034662

Miller, M. D., & Legg, S. M. (1993). Alternative assessment in a highstakes environment. *Educational Measurement: Issues and Practice*, *12*(2), 9–15. doi:10.1111/j.1745-3992.1993.tb00528.x

Miller, M. D., Linn, R. L., & Gronlund, N. E. (2009). *Measurement and assessment in teaching*. Pearson.

Mills, M. S. (2022). Promoting Inclusivity Through a Culturally Responsive Approach to Classroom Assessment Practices [Chapter]. Handbook of Research on Policies and Practices for Assessing Inclusive Teaching and Learning. IGI Global. doi:10.4018/978-1-7998-8579-5.ch018

Ministry of National Education (MEB). (2018). *Fen bilimleri dersi öğretim programı* [Science course teaching program curriculum]. Talim ve Terbiye Kurulu Başkanlığı.

Mitchelmore, S., & Rowley, J. (2010). Entrepreneurial competencies: A literature review and development agenda. *International Journal of Entrepreneurial Behaviour & Research*, *16*(2), 92–111. doi:10.1108/13552551011026995

Monbec, L., Tilakaratna, N., Brooke, M., Siew, T., Lau, Y., Shih, C., & Wu, V. (2021). *Designing a rubric for reflection in nursing: a Legitimation Code Theory and systemic functional linguistics-informed framework Designing a rubric for reflection in nursing: a Legitimation Code*. Taylor & Francis. doi:10.1080/02602938.2020.1855414

Moni, R. W., Beswick, E., & Moni, K. B. (2005). Using student feedback to construct an assessment rubric for a concept map in physiology. *Advances in Physiology Education*, *29*(4), 197–203. doi:10.1152/advan.00066.2004 PMID:16298956

Moniz, T., Arntfield, S., Miller, K., Lingard, L., Watling, C., & Regehr, G. (2015). Considerations in the use of reflective writing for student assessment: Issues of reliability and validity. *Medical Education*, *49*(9), 901–908. doi:10.1111/medu.12771 PMID:26296406

Montgomery, K. (2000). Classroom rubrics: Systematizing what teachers do naturally. *The Clearing House: A Journal of Educational Strategies, Issues and Ideas*, *73*(6), 324–328. doi:10.1080/00098650009599436

Montgomery, K. (2002). Authentic Tasks and Rubrics: Going Beyond Traditional Assessments in College Teaching. *College Teaching*, *50*(1), 34–40. doi:10.1080/87567550209595870

Moon, J. (2007). Getting the measure of reflection: Considering matters of definition and depth. *Journal of Radiotherapy in Practice*, *6*(4), 191–200. doi:10.1017/S1460396907006188

Moreno, A., Trigueros, C., & Rivera, E. (2013). Autoevaluación y Emociones en la Formación Inicial de Profesores de Educación Física. *Estudios Pedagógicos (Valdivia)*, *39*(1), 165–177. doi:10.4067/S0718-07052013000100010

Morton, J. K., Northcote, M., Kilgour, P., & Jackson, W. A. (2021). Sharing the construction of assessment rubrics with students: A Model for collaborative rubric construction. *Journal of University Teaching & Learning Practice*, *18*(4). Advance online publication. doi:10.53761/1.18.4.9

Moskal, B. M. (2003). Recommendations for developing classroom performance assessments and scoring rubrics. Practical Assessment, Research & Evaluation, 8(1), 14. ttps:// doi:10.7275/jz85-rj16

Moskal, B. M. (2000). Scoring rubrics: What, when and how? *Practical Assessment, Research & Evaluation*, *7*(3), 1–5.

Moskal, B. M., & Leydens, J. A. (2000). Scoring rubric development: Validity and reliability. *Practical Assessment, Research & Evaluation*, *7*(10), 1–6.

Moskal, B. M., & Leydens, J. A. (2000). Scoring rubric development: Validity and reliability. *Practical Assessment, Research & Evaluation*, *7*(10), 1–6. doi:10.7275/q7rm-gg74

Mouton, J. (2001). *How to succeed in your master's & doctoral studies*. Van Schaik.

Moxham, L., Dwyer, T., & Reid-Searl, K. (2013). Articulating expectations for PhD candidature upon commencement: Ensuring supervisor/student 'best fit'. *Journal of Higher Education Policy and Management*, *35*(4), 345–354. doi:10.1080/1360080X.2013.812030

Moys, J.-L., Collier, J., & Joyce, D. (2018). By design: engaging Graphic Communication students in curriculum development. *The Journal of Educational Innovation, Partnership and Change, 4* (1). https://centaur.reading.ac.uk/76158/ doi:10.21100/jeipc.v4i1.752

Mulliner, E., & Tucker, M. (2017). Feedback on feedback practice: Perceptions of students and academics. *Assessment & Evaluation in Higher Education*, *42*(2), 266–288. doi:10.1080/02602938.2015.1103365

Munezero, V., Yadav, L., & Bugingo, J. B. (2022). Representation of nature of science in physics textbooks of cycle 4 fundamental schools in Burundi. *European Journal of Educational Research*, *11*(4), 2487–2496. doi:10.12973/eu-jer.11.4.2487

Munthe, E., & Rogne, M. (2015). Research-based teacher education. *Teaching and Teacher Education*, *46*, 17–24. doi:10.1016/j.tate.2014.10.006

Murdoch, D., English, A. R., Hintz, A., & Tyson, K. (2020). *Feeling Heard*: Inclusive Education, Transformative Learning, and Productive Struggle. *Educational Theory*, *70*(5), 653–679. doi:10.1111/edth.12449

Murray, E., Treweek, S., Pope, C., MacFarlane, A., Ballini, L., Dowrick, C., Finch, T., Kennedy, A., Mair, F., O'Donnell, C., Ong, B. N., Rapley, T., Rogers, A., & May, C. (2010). Normalisation process theory: A framework for developing, evaluating and implementing complex interventions. *BMC Medicine*, *8*(1), 63. doi:10.1186/1741-7015-8-63 PMID:20961442

Muster, D. (1969). *Metodologia examinării şi notării elevilor* [Lexical characteristics of Romanian language]. E.D.P.

MYK. (2015). *Turkish Qualifications Framework*. MYK. https://www.myk.gov.tr//TRR/File6.pdf

Namal, R. (2022). Multiple examination of graduate theses within the scope of graph literacy in social studies teaching. *Karaelmas Journal of Educational Sciences*, *10*, 87–102.

National Forum for the Enhancement of Teaching and Learning. (2016). *NF Bulletin: assessment OF, FOR and AS Learning: Students as Partners in Assessment*. Teaching and Learning. https://www.teachingandlearning.ie/publication/students-as-partners/

Nespor, J. (1987). The role of beliefs in the practice of teaching. *Journal of Curriculum Studies*, *19*(4), 317–328. doi:10.1080/0022027870190403

Ng, S. L., Kinsella, E. A., Friesen, F., & Hodges, B. (2015). Reclaiming a theoretical orientation to reflection in medical education research: A critical narrative review. *Medical Education*, *49*(5), 461–475. doi:10.1111/medu.12680 PMID:25924122

Nguyen, Q. D., Fernandez, N., Karsenti, T., & Charlin, B. (2014). What is reflection? A conceptual analysis of major definitions and a proposal of a five-component model. *Medical Education*, *48*(12), 1176–1189. doi:10.1111/medu.12583 PMID:25413911

Nicol, D. (2009). Assessment for learner self-regulation: Enhancing achievement in the first year using learning technologies. *Assessment & Evaluation in Higher Education*, *34*(3), 335–352. doi:10.1080/02602930802255139

Nicol, D. J., & Macfarlane-Dick, D. (2006). Formative assessment and self-regulated learning: A model and seven principles of good feedback practice. *Studies in Higher Education*, *31*(2), 199–218. doi:10.1080/03075070600572090

Nitko, A. J., & Brookhart, S. M. (2014). *Educational Assessment of Students* (6th ed.). Pearson.

Nixon, R. S., Godfrey, T. J., Mayhew, N. T., & Wiegert, C. C. (2016). Undergraduate student construction and interpretation of graphs in physics lab activities. *Physical Review. Physics Education Research*, *12*(1), 010104. doi:10.1103/PhysRevPhysEducRes.12.010104

Nordrum, L., Evans, K., & Gustafsson, M. (2013). Comparing student learning experiences of in-text commentary and rubric-articulated feedback: Strategies for formative assessment. *Assessment & Evaluation in Higher Education*, *38*(8), 919–940. doi:10.1080/02602938.2012.758229

Norton, L. (2004). Using assessment criteria as learning criteria: A case study in psychology. *Assessment & Evaluation in Higher Education*, *29*(6), 687–702. doi:10.1080/0260293042000227236

Ntem, A., Nguyen, E., Rafferty, C., Kwan, C., & Benlahcene, A. (2020). Students as partners in crisis? Student co-editors' perspectives on COVID-19, values, and the shift to virtual spaces. *International Journal for Students as Partners*, *4*(2), 1–8. doi:10.15173/ijsap.v4i2.4432

NUS. (2015). *Assessment and feedback benchmarking tool.* QAA. https://www.qaa.ac.uk/docs/qaas/focus-on/nus-assessment-and-feedback-benchmarking-tool.pdf?sfvrsn=f37cf481_14

O'Donnell, J. A., Oakley, M., Haney, S., O'Neill, P. N., & Taylor, D. (2011). Rubrics 101: A primer for rubric development in dental education. *Journal of Dental Education*, *75*(9), 1163–1175. doi:10.1002/j.0022-0337.2011.75.9.tb05160.x PMID:21890846

O'donovan, B., Price, M., & Rust, C. (2001). The Student Experience of Criterion-Referenced Assessment (Through the Introduction of a Common Criteria Assessment Grid). *Innovations in Education and Teaching International*, *38*(1), 74–85. doi:10.1080/147032901300002873

O'donovan, B., Price, M., & Rust, C. (2004). Know what I mean? Enhancing student understanding of assessment standards and criteria. *Teaching in Higher Education*, *9*(3), 325–335. doi:10.1080/1356251042000216642

O'donovan, B., Price, M., & Rust, C. (2008). Developing student understanding of assessment standards: A nested hierarchy of approaches. *Teaching in Higher Education*, *13*(2), 205–217. doi:10.1080/13562510801923344

Oakleaf, M. (2009). *Writing rubrics right: Avoiding common mistakes in rubric assessment.* Association of College and Research Libraries 14th National Conference, Seattle, WA. http://meganoakleaf.info/writingrubricsright.pdf

Oaklef, M. (2009). Using rubrics to assess information literacy: An examination of methodology and interrater reliability. *Journal of the American Society for Information Science and Technology*, *60*(5), 969–983. doi:10.1002/asi.21030

Obschonka, M. (2016). Adolescent pathways to entrepreneurship. *Child Development Perspectives*, *10*(3), 196–201. doi:10.1111/cdep.12185

OfS. (2020). *Teaching Excellence Framework*. OfS. https://www.officeforstudents.org.uk/advice-and-guidance/teaching/about-the-tef/

OfS. (2022). *Building a culture of student engagement: our priorities for 2022-23, Student Engagement Strategy*. Available from https://www.officeforstudents.org.uk/about/student-engagement-strategy/

OfS. (2023). *Methodology document for National Student Survey sector analysis.* OfS. https://www.officeforstudents.org.uk/media/de0d31cf-23d7-42c7-8113-500aefcc96fb/nss-2022-sector-analysis-method-document.pdf

OfS. (2023). *National Student Survey (NSS)*. OfS. https://www.thestudentsurvey.com/

Compilation of References

Okojie, M. U., Bastas, M., & Fatma Miralay, F. (2022). Using Curriculum Mapping as a Tool to Match Student Learning Outcomes and Social Studies Curricula. *Frontiers in Psychology*, *13*, 850264. Advance online publication. doi:10.3389/fpsyg.2022.850264 PMID:36059751

Olesen, K. B., Christensen, M. K., & O'Neill, L. D. (2020). What do we mean by "transferable skills"? A literature review of how the concept is conceptualized in undergraduate health sciences education. *Higher Education. Skills and Work-Based Learning*, *11*(3), 616–634. doi:10.1108/HESWBL-01-2020-0012

Olson, J., & Krysiak, R. (2021). Rubrics as Tools for Effective Assessment of Student Learning and Program Quality. Curriculum Development and Online Instruction for the 21st Century (pp.173-200).

Olssen, M., & Peters, M. A. (2005). Neoliberalism, Higher Education and the Knowledge Economy: From the Free Market to Knowledge Capitalism. *Journal of Education Policy*, *20*(3), 313–345. doi:10.1080/02680930500108718

Organization for Economic Cooperation and Development (OECD). (2021). *Teaching as a Knowledge Profession: St OECD (2013) Teachers for the 21st Century: Using Evaluation to Improve Teaching*. OECD.

Orsmond, P., Merry, S., & Reiling, K. (2002). The use of formative feedback when using student derived marking criteria in peer and self-assessment. *Assessment & Evaluation in Higher Education*, *27*(4), 309–323. doi:10.1080/0260293022000001337

Örücü, D. (2012). İlköğretim sınıf öğretmenlerinin sınıfa ve sınıf yönetimine ilişkin metaforik bakışları: Karşılaştırmalı bir durum çalışması. *İlköğretim Online, 11*(2), 342-352.

Ozan, C., & Kincal, R. (2018). The effects of formative assessment on academic achievement, attitudes toward the lesson, and self-regulation skills. *Educational Sciences: Theory & Practice*, 18.

Özen, S. O. (2019). *Öğrenenlerin e-değerlendirmeye dayalı kişiselleştirilmiş geri bildirim yollarının aranması [An investigation of the learners' personalized feedback paths based on e-assessment.]* [Unpublished PhD dissertation, Eskişehir Osmangazi Univeristy, Eskişehir, Turkey].

Ozgun-Koca, S. A. (2001). The graphing skills of students in mathematics and science education. *ERIC Digest.* https://files.eric.ed.gov/fulltext/ED464804.pdf

Padden, L., & O'Neill, G. (2021). Embedding equity and inclusion in higher education assessment strategies: Creating and sustaining positive change in the post-pandemic era. In P. Baughan (Ed.), *Assessment and feedback in a post-pandemic era: A time for learning and inclusion* (pp. 138–147). Advance HE. https://www.advance-he.ac.uk/knowledge-hub/assessment-and-feedback-post-pandemic-era-time-learning-and-inclusion

Panadero, E., Broadbent, J., Boud, D., & Lodge, J. (2019). Using formative assessment to influence self- and co-regulated learning: The role of evaluative judgement. *European Journal of Psychology of Education*, *34*(3), 535–557. doi:10.100710212-018-0407-8

Panadero, E., Fraile, J., Fernández, J., Castilla-Estévez, D., & Ruíz, M. A. (2019). Spanish university assessment practices: Examination tradition with diversity by faculty. *Assessment & Evaluation in Higher Education*, *44*(3), 379–397. doi:10.1080/02602938.2018.1512553

Panadero, E., & Jonsson, A. (2013). The use of scoring rubrics for formative assessment purposes revisited. *Educational Research Review*, *9*, 129–144. doi:10.1016/j.edurev.2013.01.002

Panadero, E., & Jonsson, A. (2020). A critical review of the arguments against the use of rubrics. *Educational Research Review*, *30*, 100329. doi:10.1016/j.edurev.2020.100329

Panadero, E., Jonsson, A., & Botella, J. (2017). Effects of self-assessment on self-regulated learning and self-efficacy: Four meta-analyses. *Educational Research Review*, *22*, 74–98. doi:10.1016/j.edurev.2017.08.004

Panadero, E., & Romero, M. (2014). To rubric or not to rubric? The effects of self-assessment on self-regulation, performance and self-efficacy. *Assessment in Education: Principles, Policy & Practice*, *21*(2), 133–148. doi:10.1080/0969594X.2013.877872

Panadero, E., Romero, M., & Strijbos, J. W. (2013). The impact of a rubric and friendship on peer assessment: Effects on construct validity, performance, and perceptions of fairness and comfort. *Studies in Educational Evaluation*, *39*(4), 195–203. doi:10.1016/j.stueduc.2013.10.005

Papakonstantinou, M., & Skoumios, M. (2021). Science and engineering practices in the content of Greek middle school physics textbooks about forces and motion. *Journal of Technology and Science Education*, *11*(2), 457–473. doi:10.3926/jotse.1286

Pappas, C. C., Zecker, L. B., & Zecker, L. (2001). *Teacher inquiries in literacy teaching-learning: Learning to collaborate in elementary urban classrooms*. Routledge. doi:10.4324/9781410600769

Parker-Jenkins, M. (2018). Mind the gap: Developing the roles, expectations and boundaries in the doctoral supervisor–supervisee relationship. *Studies in Higher Education*, *43*(1), 57–71. doi:10.1080/03075079.2016.1153622

Parkes, K. A. (2006). *The effect of performance rubrics on college level applied studio grading*. [PhD dissertation. University of Maimi]. UMI No. 3215237.

Parlak, B., & Doğan, N. (2014). Dereceli puanlama anahtarı ve puanlama anahtarından elde edilen puanların uyum düzeyleri [Comparison of answer key and scoring rubric for the evaluation of the student performance]. *Hacettepe Üniversitesi Eğitim Fakültesi Dergisi / H.U. Journal of Education*, *29*(2), 189–197.

Parmar, R. S., & Signer, B. R. (2005). Sources of error in constructing and interpreting graphs a study of fourth-and fifth-grade students with LD. *Journal of Learning Disabilities*, *38*(3), 250–261. doi:10.1177/00222194050380030601 PMID:15940962

Pastore, S., & Andrade, H. L. (2019). Teacher assessment literacy: A three-dimensional model. *Teaching and Teacher Education*, *84*, 128–138. doi:10.1016/j.tate.2019.05.003

Patton, M. Q. (2008). *Utilization-focused evaluation*. Sage publications.

Peacock, S., Cowan, J., Irvine, L., & Williams, J. (2020). An Exploration into the Importance of a Sense of Belonging for Online Learners. *International Review of Research in Open and Distance Learning*, *21*(2), 18–35. https://eric.ed.gov/?id=EJ1250669. doi:10.19173/irrodl.v20i5.4539

Pegalajar, M. del C. (2021). La Rúbrica como Instrumento para la Evaluación de Trabajos Fin de Grado. REICE. *Revista Electrónica Iberoamericana sobre Calidad, Eficacia y Cambio en Educación*, *19*(3), 67–81. doi:10.15366/reice2021.19.3.005

Pekrun, R., Goetz, T., Titz, W., & Perry, R. P. (2002). Academic emotions in students' self-regulated learning and achievement: A program of qualitative and quantitative research. *Educational Psychologist*, *37*(2), 91–105. doi:10.1207/S15326985EP3702_4

Pérez-Pueyo, A., Hortigüela-Alcalá, D., & Gutiérrez-García, C. (2019). La exposición oral con pecha-kucha desde la evaluación formativa. En A. Ramírez y M. P. Gutiérrez (Coords.), La evaluación educativa, entre la emoción y la razón (pp. 104-123). Universidad de Córdoba.

Pérez-Pueyo, A., Gutiérrez-García, C., Hortigüela-Alcalá, D., & Hernando-Garijo, A. (2017). Video-diario de evidencias de aprendizaje. Infancia. *Educación y Aprendizaje*, *3*(2), 127–132. doi:10.22370/ieya.2017.3.2.711

Perkins, D. N., Jay, E., & Tishman, S. (1993). Beyond abilities: A dispositional theory of thinking. *Merrill-Palmer Quarterly*, *39*(1), 1–21.

Phage, I. B., Lemmer, M., & Hitge, M. (2017). Probing factors influencing students' graph comprehension regarding four operations in kinematics graphs. *African Journal of Research in Mathematics. Science and Technology Education*, *21*(2), 200–210.

Pilbeam, C., & Denyer, D. (2009). Lone scholar or community member? The role of student networks in doctoral education in a UK management school. [udying Pedagogical Knowledge across Education Systems. Paris: OECD Publishing.]. *Studies in Higher Education*, *34*(3), 301–318. doi:10.1080/03075070802597077

Pilcher, N., & Richards, K. (2016). The paradigmatic hearts of subjects which their 'English' flows through. *Higher Education Research & Development*, *35*(5), 997–1010. doi:10.1080/07294360.2016.1138455

Pilcher, N., & Richards, K. (2017). Challenging the power invested in the International English Language Testing System (IELTS): Why determining 'English' preparedness needs to be undertaken within the subject context. *Power and Education*, *9*(1), 3–17. doi:10.1177/1757743817691995

Pilcher, N., & Richards, K. (2022). *Enhancing Student Support in Higher Education: A Subject-focused Approach*. Springer Nature. doi:10.1007/978-3-030-81724-4

Pineda, D. (2014). The feasibility of assessing teenagers' oral English language performance with a rubric. *Profile Issues in TeachersProfessional Development*, *16*(1), 181–198. doi:10.15446/profile.v16n1.43203

Pintrich, P. R. (2002). The role of metacognitive knowledge in learning, teaching, and assessing. *Theory into Practice*, *41*(4), 219–225. doi:10.120715430421tip4104_3

Pittaway, L., & Edwards, C. (2012). Assessment: Examining practice in entrepreneurship education. *Education + Training*, *54*(8/9), 778–800. doi:10.1108/00400911211274882

Pitt, E., & Norton, L. (2017). "Now that's the feedback I want!": Students' reactions to feedback on graded work and what they do with it. *Assessment & Evaluation in Higher Education*, *42*(4), 499–516. doi:10.1080/02602938.2016.1142500

Planinic, M., Milin-Sipus, Z., Katic, H., Susac, A., & Ivanjek, L. (2012). Comparison of student understanding of line graph slope in physics and mathematics. *International Journal of Science and Mathematics Education*, *10*(6), 1393–1414. doi:10.100710763-012-9344-1

Popham, J. W. (1995). *Classroom assessment*. Allyn and Bacon.

Popham, J. W. (1997). What's wrong and what's right with rubric. *Educational Leadership*, *55*(2), 72–75.

Popham, J. W. (1997). What's wrong-and what's right-with rubrics. *Educational Leadership*, *55*(2), 72–75.

Popham, J. W. (2007). *Classroom assessment: What teachers need to know* (5th ed.). Pearson Education.

Popham, W. J. (1997). What's wrong – and what's right – with rubrics. *Educational Leadership*, *55*, 72–75.

Popham, W. J. (1997). What's Wrong—and What's Right—with Rubrics. *Educational Leadership*, *55*(2), 72–75.

Popham, W. J. (1997). What's wrong-and what's right-with rubrics? *Educational Leadership*, *55*(2), 72–75.

Popham, W. J. (2009). Assessment literacy for teachers: Faddish or fundamental? *Theory into Practice*, *48*(1), 4–11. doi:10.1080/00405840802577536

Postmes, L., Bouwmeester, R., de Kleijn, R., & van der Schaaf, M. (2021). Supervisors' untrained postgraduate rubric use for formative and summative purposes. *Assessment & Evaluation in Higher Education*, 1–14.

Prins, F. J., de Kleijn, R., & Tartwijk, J. V. (2017). Students' use of a rubric for research theses. *Assessment & Evaluation in Higher Education*, *42*(1), 128–150. doi:10.1080/02602938.2015.1085954

QAA. (2014). *UK Quality Code for Higher Education Part A: Setting and Maintaining Academic Standards PART A The Frameworks for Higher Education Qualifications of UK Degree-Awarding Bodies*. QAA. https://www.qaa.ac.uk/docs/qaa/quality-code/qualifications-frameworks.pdf

QAA. (2018). *UK Quality Code, Advice and Guidance: Student Engagement*. QAA. https://www.qaa.ac.uk/the-quality-code/advice-and-guidance/student-engagement#

Quinlan, A. M. (2012). *A complete guide to rubrics: Assessment made easy for teachers of K-college*. R&L Education.

Radu, I. T. (2008). *Evaluarea în procesul didactic* [Lexical characteristics of Romanian language]. EDP.

Ragupathi, K., & Lee, A. (2020). Beyond Fairness and Consistency in Grading: The Role of Rubrics in Higher Education. In C. S. Sanger & N. W. Gleason (Eds.), *Diversity and Inclusion in Global Higher Education: Lessons from Across Asia* (pp. 73–95). Springer. doi:10.1007/978-981-15-1628-3_3

Rahmat, R. (2017). *An Approach to Overcome the Challenges of Using Rubrics to Improve Student's Understanding of Success Criteria*. World Association of Lesson Studies.

Reddy, Y. M., & Andrade, H. (2010). A review of rubric use in higher education. *Assessment & Evaluation in Higher Education*, *35*(4), 435–448. doi:10.1080/02602930902862859

Reed, P., Watmough, S., & Duvall, P. (2015). Assessment analytics using Turnitin and Grademark in an undergraduate medical curriculum. *Journal of Perspectives in Applied Academic Practice*, *3*(2), 92–108. doi:10.14297/jpaap.v3i2.159

Reynolds-Keefer, L. (2010). Rubric-referenced assessment in teacher preparation: An opportunity to learn by using. *Practical Assessment, Research & Evaluation*, *15*(1), 8. doi:10.7275/PSK5-MF68

Rezaei, A. R., & Lovorn, M. (2010). Reliability and validity of rubrics for assessment through writing. *Assessing Writing*, *15*(1), 18–39. doi:10.1016/j.asw.2010.01.003

Rhodes, M. G. (2019). Metacognition. *Teaching of Psychology*, *46*(2), 168–175. doi:10.1177/0098628319834381

Richards, K., & Pilcher, N. (2015). Avoiding dialogues of non-discovery through promoting dialogues of discovery. *Dialogic Pedagogy, 3*.

Richards, K., & Pilcher, N. (2016). An individual subjectivist critique of the use of corpus linguistics to inform pedagogical materials. *Dialogic Pedagogy: An International Online Journal, 4*.

Richards, K., & Pilcher, N. (2017). Should we teach from materials developed with corpus linguistics? *English for Specific Purposes Special Interest Group Journal*.

Richards, K., & Pilcher, N. (2021). *Study Skills are not the answer to students woes*. WonkHE. https://wonkhe.com/blogs/study-skills-are-not-the-answer-to-students-academic-woes/ Last Accessed January 2023

Richards, K., & Pilcher, N. (2014). Contextualising higher education assessment task words with an 'anti-glossary' approach. *International Journal of Qualitative Studies in Education : QSE*, *27*(5), 604–625. doi:10.1080/09518398.2013.805443

Richards, K., & Pilcher, N. (2019). How a view of language underpins approaches to supporting higher education students that facilitate neo-liberalism, and how to resist this. *Power and Education*, *11*(1), 51–68.

Richards, K., & Pilcher, N. (2020a). Study Skills: Neoliberalism's perfect Tinkerbell. *Teaching in Higher Education*, 1–17.

Richards, K., & Pilcher, N. (2020b). Using physical objects as a portal to reveal academic subject identity and thought. *Qualitative Report*, *25*(1). Advance online publication. doi:10.46743/2160-3715/2020.4023

Robbins, J., & Marinkova, M. (in press). Students' (non)use of online rubrics: Turnitin vs feedback literacy. *Practitioner Research in Higher Education*.

Robinson, S., & Stubberud, H. A. (2014). Teaching creativity, team work and other soft skills for entrepreneurship. *Journal of Entrepreneurship Education*, *17*(2), 186.

Robley, W., Whittle, S., & Murdoch-Eaton, D. (2005a). Mapping generic skills curricula: A recommended methodology. *Journal of Further and Higher Education*, *29*(3), 221–231. doi:10.1080/03098770500166801

Robley, W., Whittle, S., & Murdoch-Eaton, D. (2005b). Mapping generic skills curricula: Outcomes and discussion. *Journal of Further and Higher Education*, *29*(4), 321–330. doi:10.1080/03098770500353342

Rogers, J., Peecksen, S., Douglas, M., & Simmons, M. (2019). *Reflective Practice International and Multidisciplinary Perspectives Validation of a reflection rubric for higher education*. Taylor & Francis. doi:10.1080/14623943.2019.1676712

Rogers, R. R. (2001). Reflection in Higher Education: A Concept Analysis. *Innovative Higher Education*, *26*(1), 37–57. doi:10.1023/A:1010986404527

Rokeach, M. (1968). *Beliefs, attitudes and values: A theory of organization and change*. Jossey-Bass Inc.

Rolfe, G., & Gardner, L. (2006). "Do not ask who I am...": Confession emancipation and (self)-management through reflection. In Journal of Nursing Management, 14 (8), 593–600. https://doi.org/ doi:10.1111/j.1365-2934.2006.00717.x

Romero, M., Castejón, F., López, V., & Fraile, A. (2017). Formative assessment, communicative competencies and ICT in teachers training. *Comunicar*, *52*, 1–12.

Rossi, P. H., Lipsey, M. W., & Freeman, H. E. (2004). *Evaluation: A Systematic Approach* (7th ed.). Sage Publications.

Ruskovaara, E., Hämäläinen, M., & Pihkala, T. (2016). HEAD teachers managing entrepreneurship education–Empirical evidence from general education. *Teaching and Teacher Education*, *55*, 155–164. doi:10.1016/j.tate.2016.01.004

Ruskovaara, E., & Pihkala, T. (2013). Teachers implementing entrepreneurship education: Classroom practices. *Education + Training*, *55*(2), 204–216. doi:10.1108/00400911311304832

Rust, C., Price, M., & O'Donovan, B. (2003). Improving Students' Learning by Developing Their Understanding of Assessment Criteria and Processes. *Assessment & Evaluation in Higher Education*, *28*(2), 147–164. doi:10.1080/02602930301671

Ryan, J., & Viete, R. (2009). Respectful interactions: Learning with international students in the English-speaking academy. *Teaching in Higher Education*, *14*(3), 303–314. doi:10.1080/13562510902898866

Ryan, R. M., & Deci, E. L. (2000). Intrinsic and extrinsic motivations: Classic definitions and new directions. *Contemporary Educational Psychology*, *25*(1), 54–67. doi:10.1006/ceps.1999.1020 PMID:10620381

Saddler, B., & Andrade, H. (2004). The writing rubric. *Educational Leadership*, *62*(2), 48–52.

Sadler, D. R. (1983). Evaluation and the improvement of academic learning. *The Journal of Higher Education*, *54*(1), 60–79. doi:10.2307/1981645

Sadler, D. R. (1989). Formative assessment and the design of instructional systems. *Instructional Science*, *18*(2), 119–144. doi:10.1007/BF00117714

Sadler, D. R. (1998). Formative assessment: Revisiting the territory. *Assessment in Education: Principles, Policy & Practice*, *5*(1), 77–84. doi:10.1080/0969595980050104

Sadler, D. R. (2005). Interpretations of criteria-based assessment and grading in higher education. *Assessment & Evaluation in Higher Education*, *30*(2), 175–194. doi:10.1080/0260293042000264262

Sadler, D. R. (2009). *Transforming holistic assessment and grading into a vehicle for complex learning*. Springer. doi:10.1007/978-1-4020-8905-3_4

Sadler, D. R. (2009a). Transforming holistic assessment and grading into a vehicle for complex learning. *Assessment, Learning and Judgement in Higher Education*. In G. Joughin (Ed.), *Springer Science+Business Media B.V.*, doi:10.1007/978 1 4020 8905 3_4

Sadler, D. R. (2009b). Indeterminacy in the use of preset criteria for assessment and grading in higher education. *Assessment & Evaluation in Higher Education*, *34*(2), 159–179. doi:10.1080/02602930801956059

Sadler, D. R. (2010). Beyond feedback: Developing student capability in complex appraisal. *Assessment & Evaluation in Higher Education*, *35*(5), 535–550. doi:10.1080/02602930903541015

Sadler, D. R. (2011). Academic freedom, achievement standards and professional identity. *Quality in Higher Education*, *17*(1), 85–100. doi:10.1080/13538322.2011.554639

Sadler, D. R. (2013). Assuring academic achievement standards: From moderation to calibration. *Assessment in Education: Principles, Policy & Practice*, *20*(1), 5–19. doi:10.1080/0969594X.2012.714742

Sadler, R. D. (2014). The futility of attempting to codify academic achievement standards. *Higher Education*, *67*(3), 273–288. doi:10.100710734-013-9649-1

Salazar-Torres, J., Leal, O. R., & Ortega, M. V. (2021). The rubric as an assessment tool for solving problem situations in the physics and mathematics teaching context. *Journal of Physics: Conference Series: V International Meeting of Mathematical Education*. IOP Science. 10.1088/1742-6596/1981/1/012018

Sambell, K., & McDowell, L. (1998). The construction of the hidden curriculum: Messages and meanings in the assessment of student learning. *Assessment & Evaluation in Higher Education*, *23*(4), 391–402. doi:10.1080/0260293980230406

Santos Guerra, M. A. (2003). Dime cómo evalúas y te diré qué tipo de profesional y de persona eres. *Revista Enfoques Educacionales*, *5*(1), 1–15. doi:10.5354/0717-3229.2003.47513

Santos Guerra, M. A. (2014). *La evaluación como aprendizaje. Cuando la flecha impacta en la diana*. Narcea.

Sarja, A., & Janhonen, S. (2009). Methodological reflections: Supervisory discourses and practice-based learning. *Teaching in Higher Education*, *14*(6), 619–630. doi:10.1080/13562510903315100

Sawyer, R., & Norris, J. (2015). Duoethnography. *International Journal of Qualitative Research*, *8*(1), 1–4. doi:10.1525/irqr.2015.8.1.1

Schelfhout, W., Bruggemana, K., & De Maeyerb, S. (2016). Evaluation of entrepreneurial competence through scaled behavioural indicators: Validation of an instrument. *Studies in Educational Evaluation*, *51*, 29–41. doi:10.1016/j.stueduc.2016.09.001

Schleppegrell, M. J. (2004). *The language of schooling: A functional linguistics perspective*. Routledge. doi:10.4324/9781410610317

Schmidt, M. & Hansson, E., (2021). "I didn't want to be a troublemaker"–doctoral students' experiences of change in supervisory arrangements. *Studies in Graduate and Postdoctoral Education*.

Scholz, U., Doña, B. G., Sud, S., & Schwarzer, R. (2002). Is general self-efficacy a universal construct? Psychometric findings from 25 countries. *European Journal of Psychological Assessment*, *18*(3), 242. doi:10.1027//1015-5759.18.3.242

Schön, D. A. (1983). *The reflective practitioner : how professionals think in action*. Taylor & Francis.

Schraw, G. (1998). Promoting general metacognitive awareness. *Instructional Science*, *26*(1/2), 113–125. doi:10.1023/A:1003044231033

Schreiber, L. M., Paul, G. D., & Shibley, L. R. (2012). The development and test of the public speaking competence rubric. *Communication Education*, *61*(3), 205–233. doi:10.1080/03634523.2012.670709

Schumpeter, J. A. (1982). *The theory of economic development: An inquiry into profits, capital, credit, interest, and the business cycle. Harward Economic Studies 46*. Harvard University Press.

Scott, J. C. (1990). *Domination and the arts of resistance: Hidden transcripts*. Yale university press.

Scriven, M. (1991). *Evaluation Thesaurus* (4th ed.). Sage Publications.

Seçken, N., & Çelik, Ç. (2021). Investigating high school students' graphic interpretation skills on the subject of chemical equilibrium. *Journal of Research in Education and Society*, *8*(1), 179–204.

Seikkula-Leino, J. (2011). The implementation of entrepreneurship education through curriculum reform in Finnish comprehensive schools. *Journal of Curriculum Studies*, *43*(1), 69–85. doi:10.1080/00220270903544685

Shabani, E. A., & Panahi, J. (2020). Examining consistency among different rubrics for assessing writing. *Language Testing in Asia*, *10*(1), 1–25. doi:10.118640468-020-00111-4

Short, J. C., Ketchen, D. J. Jr, Shook, C. L., & Ireland, R. D. (2010). The concept of "opportunity" in entrepreneurship research: Past accomplishments and future challenges. *Journal of Management*, *36*(1), 40–65. doi:10.1177/0149206309342746

Silveyra, G., Herrero, Á., & Pérez, A. (2021). Model of teachable entrepreneurship competencies (M-TEC): Scale development. *International Journal of Management Education*, *19*(1), 100392. doi:10.1016/j.ijme.2020.100392

Skillings, M. J., & Ferrell, R. (2000). Student-Generated Rubrics: Bringing Students into the Assessment Process. *The Reading Teacher*, *53*(6), 452–455.

Slutsky, D. J. (2014). The effective use of graphs. *Journal of Wrist Surgery, 03*(02), 067-068. doi:10.1055/s-0034-1375704

Smith, M., & Trede, F. (2013). *Higher Education Research & Development Reflective practice in the transition phase from university student to novice graduate: implications for teaching reflective practice*. Taylor & Francis Online. doi:10.1080/07294360.2012.709226

Smith, C. D., Worsfold, K., Davies, L., Fisher, R., & McPhail, R. (2013). Assessment literacy and student learning: The case for explicitly developing students 'assessment literacy'. *Assessment & Evaluation in Higher Education*, *38*(1), 44–60. doi:10.1080/02602938.2011.598636

Smit, R., Bachmann, P., Blum, V., Birri, T., & Hess, K. (2017). Effects of a rubric for mathematical reasoning on teaching and learning in primary school. *Instructional Science*, *45*(5), 603–622. doi:10.100711251-017-9416-2

Snelling, C. A., Loveys, B. R., Karanicolas, S., Schofield, N. J., Carlson-Jones, W., Weissgerber, J., Edmonds, R., & Ngu, J. (2019). Partnership through co-creation: Lessons learnt at the University of Adelaide. *International Journal for Students as Partners*, *3*(2), 2. doi:10.15173/ijsap.v3i2.3799

Soemantri, D., Mustika, R., & Greviana, N. (2022). Inter-Rater Reliability of Reflective-Writing Assessment in an Undergraduate Professionalism Course in Medical Education. *Education in Medicine Journal, 14*(1), 87–97.

Sommarström, K., Oikkonen, E., & Pihkala, T. (2021). The school and the teacher autonomy in the implementing process of entrepreneurship education curricula. *Education Sciences*, *11*(5), 215. doi:10.3390/educsci11050215

Spencer, D., Riddle, M., & Knewstubb, B. (2012). Curriculum mapping to embed graduate capabilities. *Higher Education Research & Development*, *31*(2), 217–231. doi:10.1080/07294360.2011.554387

Stanley, T. (2021). *Using rubrics for performance-based assessment: A practical guide to evaluating student work*. Routledge. doi:10.4324/9781003239390

Stassen, M. L. A., Doherty, K., & Poe, M. (2011). *Course-based review and assessment: methods for understanding student learning*. Robert Langhorst & Company Booksellers.

Stellmack, M. A., Konheim-Kalkstein, Y. L., Manor, J. E., Massey, A. R., & Schmitz, J. A. P. (2009). An assessment of reliability and validity of a rubric for grading APA-style introductions. *Teaching of Psychology*, *36*(2), 102–107. doi:10.1080/00986280902739776

Stemler, S. E. (2004). A comparison of consensus, consistency and measurement approaches to estimating interrater reliability. *Practical Assessment, Research & Evaluation*, 9.

Stenhouse, L. (1975). *An introduction to curriculum research and development*. Heinemann.

Stenhouse, L. (1985a). What counts as research. In J. Rudduck & D. Hopkins (Eds.), *Research as a basis for teaching: Readings from the work of Lawrence Stenhouse*. Heinemann Educational Books.

Stenhouse, L. (1985b). The objectives model: some limitations. In J. Rudduck & D. Hopkins (Eds.), *Research as a basis for teaching: Readings from the work of Lawrence Stenhouse*. Heinemann Educational Books.

Stevens, D.D. & Levi, A. J. (2005). Leveling the field: Using rubrics to achieve greater equity in teaching and assessment. *Essays on Teaching Excellence, Professional and Organizational Development Network in Higher Education, 17* (1).

Stevens, D. D., & Levi, A. (2013). *Introduction to rubrics: an assessment tool to save grading time, convey effective feedback, and promote student learning* (2nd ed.). Sterling.

Stevens, D. D., & Levi, A. J. (2004). Introduction to Rubrics: An Assessment Tool to Save Grading Time, Convey Effective Feedback and Promote Student Learning. In *Stylus Publishing, LLC*. Stylus Publishing, LLC.

Stevens, D. D., & Levi, A. J. (2013). *Introduction to rubrics: An assessment tool to save grading time, convey effective feedback, and promote student learning*. Stylus Publishing, LLC.

Stiggins, R. J. (2001). *Student-involved classroom assessment* (3rd ed.). Prentice-Hall.

Stoica, A. (2003). *Evaluarea progresului scolar: de la teorie la practica* [Lexical characteristics of Romanian language]. Humanitas Educational.

Student Minds. (2019). *The University Mental Health Charter Framework*. University Mental Health Charter. https://universitymentalhealthcharter.org.uk/themes/

Stupans, I., March, G., & Owen, S. M. (2012). *Assessment & Evaluation in Higher Education Enhancing learning in clinical placements: reflective practice, self-assessment, rubrics and scaffolding*. Taylor & Francis Online. doi:10.1080/02602938.2012.658017

Sumsion, J., & Fleet, A. (1996). Reflection: Can we assess it? Should we assess it? *Assessment & Evaluation in Higher Education*, *21*(2), 121–130. doi:10.1080/0260293960210202

Sumsion, J., & Goodfellow, J. (2004). Identifying generic skills through curriculum mapping: A critical evaluation. *Higher Education Research & Development*, *23*(3), 329–346. doi:10.1080/0729436042000235436

Suskie, L. (2017). Rubric Development. In C. Secolsky & D. B. Denison (Eds.), *Handbook on measurement, assessment, and evaluation in Higher Education* (2nd ed., pp. 545–559). Routledge. doi:10.4324/9781315709307-43

Sutton, P. (2012). Conceptualizing feedback literacy: Knowing, being, and acting. *Innovations in Education and Teaching International*, *49*(1), 31–40. doi:10.1080/14703297.2012.647781

Sverdlik, A., Hall, N. C., McAlpine, L., & Hubbard, K. (2018). The PhD experience: A review of the factors influencing doctoral students' completion, achievement, and well-being. *International Journal of Doctoral Studies*, *13*, 361–388. doi:10.28945/4113

Swales, J. (1990). *'Genre Analysis: English in Academic and Research Settings*. Cambridge University Press.

Tafel, L. S., & Fischer, J. C. (1996). Lives of inquiry: Communities of learning and caring. In G. Burnaford, J. Fischer, & D. Hobson (Eds.), *Teachers doing research: Practical possibilities. Mahwah, N.J. Teaching Council of Ireland. (2017). Initial Teacher Education: Criteria and Guidelines for Programme Providers*. Teaching Council.

Tahahashi, K., & Kiyosumi, M. (2021). Development of Rubric for enhancing Sensemaking among Team members in PBL in Entrepreneurship Education. In *International Symposium on Affective Science and Engineering ISASE2021*. Japan Society of Kansei Engineering. 10.5057/isase.2021-C000004

Tairab, H. H., & Khalaf-AlNaqbi, A. K. (2004). How do secondary school science students interpret and construct scientific graphs? *Journal of Biological Education*, *38*(3), 127–132. doi:10.1080/00219266.2004.9655920

Tan, K. H. K. (2020). *Assessment Rubrics Decoded*. Routledge. doi:10.4324/9780429022081

Tapprich, W. E., Reichart, L., Simon, D. M., Duncan, G., McClung, W., Grandgenett, N., & Pauley, M. A. (2021). An instructional definition and assessment rubric for bioinformatics instruction. *Biochemistry and Molecular Biology Education*, *49*(1), 38–45. doi:10.1002/bmb.21361 PMID:32744803

Tarakçı, F. (2016). *Examining science teacher candidates' abilities on reading, interpreting and preparing of graphs.* [Unpublished master's thesis, Kastamonu University, Türkiye].

Tarnoff, K. (2023). How to Make Sure an AoL System Is Working. https://www.aacsb.edu/insights/articles/2023/06/how-to-make-sure-an-aol-system-is-working

Taşar, M. F., İngeç, Ş. K., & Güneş, P. Ü. (2002). *Determining graphic drawing and understanding skills.* V. National Science and Mathematics Education Congress, ODTÜ, Ankara.

Teaching Council of Ireland. (2020). *Céim: Standards in Initial Teacher Education in the Republic of Ireland*. Teaching Council.

Tenam-Zemach, M., & Flynn, J. E. (2015). *Rubric nation, critical inquiries on the impact of rubrics in education*. Information Age Publishing, Inc.

Teodorescu, R. E., Bennhold, C., Feldman, G., & Medsker, L. (2014). Curricular reforms that improve students' attitudes and problem-solving performance. *European Journal of Physics Education*, *5*(1), 15–44. doi:10.20308/ejpe.91287

Texas A&M Commerce University (2016). *Rubrics for assessing dissertations*. Texas A & M Commerce University. https://www.tamuc.edu/academics/graduateschool/Thesis%20and%20Dissertation%20Services/graduate%20rep%20rubrics.aspx

TGER (2016). Turkish Graduate Education Regulation. *27561 numbered Official Gazette*.

Tierney, R. D. (2013). Fairness in classroom assessment. In J. H. McMillan (Ed.), *SAGE handbook of research on classroom assessment* (pp. 125–144). SAGE Publications. doi:10.4135/9781452218649.n8

Tierney, R. D. (2022). Fairness in educational testing and assessment. In D. Fisher (Ed.), *Routledge encyclopedia of education*. Routledge. doi:10.4324/9781138609877-REE35-1

Tierney, R., & Simon, M. (2004). What's still wrong with rubrics: Focusing on the consistency of performance criteria across scale levels. *Practical Assessment, Research & Evaluation*, *9*(2), 1–7.

Topping, K. J. (2009). Peer assessment. *Theory into Practice*, *48*(1), 20–27. doi:10.1080/00405840802577569

Törner, G. (2002). Mathematical beliefs-A search for a common ground: Some theoretical considerations on structuring beliefs, some research questions, and some phenomenological observations. In G. Leder, E. Pehkonen, & G. Törner (Eds.), *Beliefs: A Hidden Variable in Mathematics Education?* (pp. 73–94). Kluwer. doi:10.1007/0-306-47958-3_5

Torrance, H. (2007). Assessment as Learning? How the Use of Explicit Learning Objectives, Assessment Criteria and Feedback in Post-Secondary Education and Training can come to Dominate Learning. *Assessment in Education: Principles, Policy & Practice*, *14*(3), 281–294. doi:10.1080/09695940701591867

Tortorella, G., Viana, S., & Fettermann, D. (2015). Learning cycles and focus groups: A complementary approach to the A3 thinking methodology. *The Learning Organization*, *22*(4), 229–240. doi:10.1108/TLO-02-2015-0008

Troyan, F. J., Sembiante, S. F., & King, N. (2019). A case for a functional linguistic knowledge base in world language teacher education. *Foreign Language Annals*, *52*(3), 644–669. doi:10.1111/flan.12410

Tsakiridou, H., & Stergiou, K. (2014). Entrepreneurial competences and entrepreneurial intentions of students in primary education. *International Journal of Humanities Social Sciences and Education*, *1*(9), 106–117.

Turley, E. D., & Gallagher, C. W. (2008). On the" uses" of rubrics: Reframing the great rubric debate. *English Journal*, 87–92.

Turnitin. (2021). *A New Path and Purpose for Turnitin*. Turnitin. https://www.turnitin.com/blog/a-new-path-and-purpose-for-turnitin

Ulker, V. (2017). The Design and Use of Speaking Assessment Rubrics. *Journal of Education and Practice*.

Unin, N., & Bearing, P. (2016). Brainstorming as a Way to Approach Student-centered Learning in the ESL Classroom. *Procedia: Social and Behavioral Sciences*, *224*, 605–612. doi:10.1016/j.sbspro.2016.05.450

University of Reading. (2018). *Marking and Feedback*. University of Reading. https://www.reading.ac.uk/cqsd/-/media/project/functions/cqsd/documents/qap/10-marking-withannexes.pdf?la=en&hash=AF70ED315F16D1ACC0322306C53138DE

University of Reading. (2020). *Partnership at UoR Guide*. University of Reading. https://sites.reading.ac.uk/wp-content/uploads/sites/35/2020/10/Partnership-at-UoR-Guide-October-2020.pdf

University of Reading. (2021). *Curriculum Framework*. University of Reading. https://www.reading.ac.uk/cqsd/-/media/project/functions/cqsd/documents/qap/university-of-reading-curriculum-framework.pdf

Ustav, S., & Venesaar, U. (2018). Bridging metacompetencies and entrepreneurship education. *Education + Training*, *60*(7/8), 674–695. doi:10.1108/ET-08-2017-0117

Uygur, J., Stuart, E., De Paor, M., Wallace, E., Duffy, S., O'Shea, M., Smith, S., & Pawlikowska, T. (2019). A Best Evidence in Medical Education systematic review to determine the most effective teaching methods that develop reflection in medical students: BEME Guide No. 51. *Medical Teacher*, *41*(1), 3–16. doi:10.1080/0142159X.2018.1505037 PMID:30634872

Valle, C., Andrade, H., Palma, M., & Hefferen, J. (2016). Applications of peer and self-assessment in music education. *Music Educators Journal*, *102*(4), 41–49. doi:10.1177/0027432116644652

Värlander, S. (2008). The Role of Students' Emotions in Formal Feedback Situations. *Teaching in Higher Education 13* (2), 145–156.

Vasileiadou, D., & Karadimitriou, K. (2021). Examining the impact of self-assessment with the use of rubrics on primary school students' performance. *International Journal of Educational Research Open*, *2*, 100031. doi:10.1016/j.ijedro.2021.100031

Vatansever, K. (2011). Öğretim üyelerinin öğretme yönelimlerinin belirlenmesi: Klinisyen öğretim üyeleriyle niteliksel bir çalışma (Identifying clinical teachers' orientation towards teaching: Qualitative survey of clinical teachers). *Tıp Eğitimi Dünyası*, *29*, 34–47.

Velasco-Martínez, L. C., & Tójar-Hurtado, J. C. (2018). Competency-Based Evaluation in Higher Education—Design and Use of Competence Rubrics by University Educators. *International Education Studies*, *11*(2), 118–132. doi:10.5539/ies.v11n2p118

Venesaar, U., Malleus, E., Arro, G., & Toding, M. (2021). Entrepreneurship Competence Model for Supporting Learners Development at All Educational Levels. *Administrative Sciences*, *12*(1), 2. doi:10.3390/admsci12010002

Verschuren, P. (2003). Case study as a research strategy: Some ambiguities and opportunities. *International Journal of Social Research Methodology*, *6*(2), 121–139. doi:10.1080/13645570110106154

Vieira, I. (2013). *A autoavaliação como instrumento de regulação da aprendizagem*. [Dissertação apresentada à Universidade Aberta]

Voloshinov, V. N., & Titurin, I. R. (1987). Freudianism: A Critical Sketch. New York: Seminar Press [original 1927] In Morris, P (Ed.) The Bakhtin Reader: Selected Writings of Bakhtin, Medvedev, Voloshinov.

Voloshinov, V. N., Matejka, L., & Titunik, I. R. (1973). *Marxism and the philosophy of language*. Seminar Press. [original 1929]

Voltmer, J.-B., Reich-Stiebert, N., Raimann, J., & Stürmer, S. (2022). The role of multi-attributional student diversity in computer-supported collaborative learning. *The Internet and Higher Education*, *55*, 100868. doi:10.1016/j.iheduc.2022.100868

Vučeljić, M., & Šuškačević, M. (2016). Achievements of Montenegrin high-school students in TUGK test (test of understanding graphs – kinematics). *AIP Conference Proceedings*, *1722*, 310007–3. doi:10.1063/1.4944317

Wald, H. S. (2015). Refining a definition of reflection for the being as well as doing the work of a physician. *Medical Teacher*, *37*(7), 696–699. doi:10.3109/0142159X.2015.1029897 PMID:25897706

Wald, H. S., Borkan, J. M., Taylor, J. S., Anthony, D., & Reis, S. P. (2012). Fostering and evaluating reflective capacity in medical education: Developing the REFLECT rubric for assessing reflective writing. *Academic Medicine*, *87*(1), 41–50. doi:10.1097/ACM.0b013e31823b55fa PMID:22104060

Walker, S., & Hobson, J. (2014). Interventions in teaching first-year law: Feeding forward to improve learning outcomes. *Assessment & Evaluation in Higher Education*, *39*(3), 326–338. doi:10.1080/02602938.2013.832728

Walland, E., & Shaw, S. (2022). E-portfolios in teaching, learning and assessment: Tensions in theory and praxis. *Technology, Pedagogy and Education*, *31*(3), 363–379. doi:10.1080/1475939X.2022.2074087

Wang, W. (2017). Using rubrics in student self-assessment: Student perceptions in the English as a foreign language writing context. *Assessment & Evaluation in Higher Education*, *42*(8), 1280–1292. doi:10.1080/02602938.2016.1261993

Wang, Z. H., Wei, S., Ding, W., Chen, X., Wang, X., & Hu, K. (2012). Students cognitive reasoning of graphs: Characteristics and progression. *International Journal of Science Education*, *34*(13), 2015–2041. doi:10.1080/09500693.2012.709333

Wanzare, Z. (2007). The transition process: The early years of being a teacher. In *Handbook of Teacher Education*. Springer., doi:10.1007/1-4020-4773-8_23

Watts, L. (2015). An Autoethnographic Exploration of Learning and Teaching Reflective Practice. *Social Work Education*, *34*(4), 363–376. doi:10.1080/02615479.2015.1016903

Watty, K., Freeman, M., Howieson, B., Hancock, P., O'Connell, B., de Lange, P., & Abraham, A. (2014). Social moderation, assessment and assuring standards for accounting graduates. *Assessment & Evaluation in Higher Education*, *39*(4), 461–478. doi:10.1080/02602938.2013.848336

Weigle, S. C. (2002). *Assessing writing*. Cambridge University Press. doi:10.1017/CBO9780511732997

Weldmeskel, F. M., & Michael, D. J. (2016). The impact of formative assessment on self-regulating learning in university classrooms. *Tuning Journal for Higher Education*, *4*(1), 99–118. doi:10.18543/tjhe-4(1)-2016pp99-118

Wenden, A. (1987). Metacognition: An expanded view of the cognitive abilities of L2 learners. *Language Learning*, *37*(4), 573–594. doi:10.1111/j.1467-1770.1987.tb00585.x

Wiggins, G. (1990). The Case for Authentic Assessment. *Practical Assessment, Research & Evaluation*, *2*(1), 2. doi:10.7275/FFB1-MM19

Wiggins, G. (1998). *Educative sssessment. Designing assessments to inform and improve student performance*. Jossey-Bass.

Wiliam, D. (2000). *Integrating Summative And Formative Functions Of Assessment*.

Willis, J., Adie, L., & Klenowski, V. (2013). Conceptualizing teachers' assessment literacies in an era of curriculum and assessment reform. *Australian Educational Researcher, 40*(2), 241e256

Willis, J., Adie, L., & Klenowski, V. (2013). Conceptualising teachers' assessment literacies in an era of curriculum and assessment reform. *Australian Educational Researcher*, *40*(2), 241–256. doi:10.100713384-013-0089-9

Wilson, M. (2007). Why I won't be using rubrics to respond to students' writing. *English Journal*, *96*(4), 62–66. doi:10.2307/30047167

Wingate, U. (2015). *Academic literacy and student diversity*. Multilingual Matters. doi:10.21832/9781783093496

Winkel, A. F., Yingling, S., Jones, A.-A., & Nicholson, J. (2017). Reflection as a Learning Tool in Graduate Medical Education: A Systematic Review. *Journal of Graduate Medical Education*, *9*(4), 430–439. doi:10.4300/JGME-D-16-00500.1 PMID:28824754

Winsor, P. J., & Ellefson, B. A. (1995). Professional portfolios in teacher education: An exploration of their value and potential. [Taylor & Francis Online], [Google Scholar]. *Teacher Educator*, *31*(1), 68–81. doi:10.1080/08878739509555100

Winstone, N. E., Nash, R. A., Parker, M., & Rowntree, J. (2017). Supporting Learners' Agentic Engagement With Feedback: A Systematic Review and a Taxonomy of Recipience Processes. *Educational Psychologist*, *52*(1), 17–37. doi:10.1080/00461520.2016.1207538

Winstone, N., & Carless, D. (2022). *Designing effective feedback processes in Higher Education: a learning-focused approach*. Routledge.

Wolcott, W., & Legg, S. M. (1998) *An Overview of Writing Assessment: Theory, Research, and Practice*. National Council of Teachers of English. https://files.eric.ed.gov/fulltext/ED423541.pdf

Wolf, K., & Stevens, E. (2007). The role of rubrics in advancing and assessing student learning. *The Journal of Effective Teaching*, *7*(1), 3–14.

Wollenschläger, M., Hattie, J., Machts, N., Möller, J., & Harms, U. (2016). What makes rubrics effective in teacher-feedback? Transparency of learning goals is not enough. *Contemporary Educational Psychology*, *44*, 1–11. doi:10.1016/j.cedpsych.2015.11.003

Wood, P. (2017). Overcoming the problem of embedding change in educational organizations: A perspective from Normalization Process Theory. *Management in Education*, *31*(1), 33–38. doi:10.1177/0892020616685286

Woodruffe, C. (1993). What Is Meant by a Competency? *Leadership and Organization Development Journal*, *14*(1), 29–36. doi:10.1108/eb053651

Wright, K., Slaboch, P. E., & Jamshidi, R. (2022). Technical writing improvements through engineering lab courses. *International Journal of Mechanical Engineering Education*, *50*(1), 120–134. doi:10.1177/0306419020939621

Wubbels, T., & Levy, J. (1993). *Interpersonal relationships in education*. The Falmer Press.

Yabanlı, H., Yıldırım, B., & Günaydın, Ö. (2013). Translating from map to line graph. [ATED]. *Araştırma Temelli Etkinlik Dergisi*, *3*(1), 12–19.

Yang, M., & Carless, D. (2013). The feedback triangle and the enhancement of dialogic feedback processes. *Teaching in Higher Education*, *18*(3), 285–297. doi:10.1080/13562517.2012.719154

Yan, Z. (2019). Self-assessment in the process of self-regulated learning and its relationship with academic achievement. *Assessment & Evaluation in Higher Education*, *45*(2), 224–238. doi:10.1080/02602938.2019.1629390

Yan, Z., Wang, X., Boud, D., & Lao, H. (2021). The effect of self-assessment on academic performance and the role of explicitness: A meta-analysis. *Assessment & Evaluation in Higher Education*. Advance online publication. doi:10.1080/02602938.2021.2012644

Yayla, G., & Özsevgeç, T. (2014). The examination of secondary school students' graphic skills: Construction and interpretation of line graphs. K. Ü. *Kastamonu Education Journal*, *23*(3), 1381–1400.

Yeltekin Atar, B. Ş., & Aykutlu, I. (2023). High school students' user skills concerning force and motion graphs. *Gazi University Journal of Gazi Education Faculty*, *43*(1), 211–242. doi:10.17152/gefad.1205369

Yılmaz, K. (2021). Systematic review, meta evaluation, and bibliometric analysis in social sciences and educational sciences. *MANAS Journal of Social Studies*, *10*(2), 1457–1490.

Yin, R. K. (2012). Case study methods. In H. Cooper, P. M. Camic, D. L. Long, A. T. Panter, D. Rindskopf, & K. J. Sher (pp. 141–155.), APA handbook of research methods in psychology, Vol. 2. Research designs: Quantitative, qualitative, neuropsychological, and biological. American Psychological Association. doi:10.1037/13620-009

Young, P. (2000). "I might as well give up": Self-esteem and mature students' feelings about feedback on assignments. *Journal of Further and Higher Education*, *24*(3), 409–418. doi:10.1080/030987700750022325

Yurdugül, H. (2005). *Ölçek Geliştirme Çalışmalarında Kapsam Geçerliği için Kapsam Geçerlik İndekslerinin Kullanılması.* Paper presented at XIV Ulusal Eğitim Bilimleri Kongresi, Pamukkale Üniversitesi Eğitim Fakültesi, Denizli. http://yunus.hacettepe.edu.tr/~yurdugul/3/indir/PamukkaleBildiri.pdf

Zeichner, K. M., & Noffke, S. E. (2001). Practitioner research. In V. Richardson (Ed.), Handbook of research on teaching (4th ed., pp. 298–330).

Zhang, Y., Chen, B. L., Ge, J., Hung, C. Y., & Mei, L. (2019). When is the best time to use rubrics in flipped learning? A study on students' learning achievement, metacognitive awareness, and cognitive load. *Interactive Learning Environments*, *27*(8), 1207–1221. doi:10.1080/10494820.2018.1553187

Zhao, C. M., Golde, C. M., & McCormick, A. C. (2007). More than a signature: How advisor choice and advisor behaviour affect doctoral student satisfaction. *Journal of Further and Higher Education*, *31*(3), 263–281. doi:10.1080/03098770701424983

Zhao, X., Cox, A., Lu, A., & Alsuhaibani, A. (2022). A comparison of student and staff perceptions and feelings about assessment and feedback using cartoon annotation. *Journal of Further and Higher Education*, *46*(5), 586–604. doi:10.1080/0309877X.2021.1986620

Zimmaro, D. M. (2004). *Developing grading rubrics*. BSU. http://bsuenglish101.pbworks.com/f/rubricshandout.pdf

Zimmerman, B. J. (1990). Self-regulated learning and academic achievement: An overview. *Educational Psychologist*, *25*(1), 3–17. doi:10.120715326985ep2501_2

Zimmerman, B. J. (2000). Attaining Self-Regulation: A Social Cognitive Perspective. In M. Boekaerts, P. R. Pintrich, & M. Zeidner (Eds.), *Handbook of Self-Regulation* (pp. 13–39). Academic Press., doi:10.1016/B978-012109890-2/50031-7

Zimmermann, P., Flavier, É., & Méard, J. (2012). L'identité professionnelle des enseignants en formation initiale. *Spiral-E. Revue de recherches en éducationı, 49*, 35-50. https://doi.org/ doi:10.3406/spira.2012.1724

About the Contributors

Chahna Gonsalves is a Lecturer in Marketing (Education) at King's Business School. She has several years of experience teaching modules at the undergraduate and postgraduate levels, including Principles of Marketing, Brand Management, Advertising, Integrated and Digital Marketing Communications and Digital Marketing. Chahna holds a PhD in Marketing and an MA in International Business & Management from the University of Westminster and obtained her BA in International Management and Business Administration with French from the University of Reading. Prior to joining King's, Chahna was a Lecturer at Westminster Business School and a Teaching Associate at Surrey Business School. Chahna is the Education Lead for the Marketing Group at King's College London. She won the 2023 King's College London Dean's Award for outstanding contribution to education. She is a Senior Fellow of the Higher Education Academy (SFHEA), Deputy Chair of the Academy of Marketing Education Special Interest Group and a Chartered Business Management Educator (CMBE).

Jayne Pearson is a Lecturer in Education with 15 years' experience in teaching in national and international contexts before joining King's Academy in 2018 as an educational developer. She is the faculty liaisons for King's Business School, the IoPPN and King's Foundations. As part of the KA team, she designs and delivers pedagogic support in many areas for academic and professional services staff in line with Kings Education Strategy. Her main interest and areas of expertise are assessment and feedback, and much of her research is in this area, and she is currently the College Academic Lead for assessment and feedback and chair of the College Assessment Working Group.

<center>***</center>

Catarina Amorim is a Sport Science PhD candidate at University of Coimbra researching topics related to professional identity formation of Physical Education teachers. She is especially interested in how professional identity is formed and developed during initial teacher education.

Heidi Andrade, Ed.D., is a Professor of Educational Psychology and Methodology at the University at Albany, SUNY. Her work focuses on the relationships between learning and assessment, with emphases on student self-assessment and self-regulated learning. She has written many articles, including an early, award-winning article on rubrics for Educational Leadership (1996). She has authored, edited and co-edited several books on classroom assessment, including the Handbook of Formative Assessment in the Disciplines (2019), Using Assessment to Enhance Learning, Achievement, and Academic Self-Regulation (2017), the SAGE Handbook of Research on Classroom Assessment (2013), and the

Handbook of Formative Assessment (2010). She has edited or co-edited special issues of Theory Into Practice (2009) and Applied Measurement in Education (2013). A long-term working relationship with arts educators in New York City has produced a collection of formative assessments for the arts: . Her current research involves the use of the Diagnostic Assessment and Achievement of College Skills () to support newly enrolled college students.

Işıl Aykutlu is an associate professor at Hacettepe University, Department of Physics Education. She is interested in different topics in physics education, such as concept maps, conceptual understanding, curriculum, metaphors, and students' ability to use graphics.

Katharine Babbitt is currently a postgraduate student in the Master of Education program at Mary Immaculate College in Limerick, Ireland. Before moving to Ireland, Katharine received her undergraduate degree and teaching credential from California State University of Long Beach in Long Beach, California. Katharine is currently writing her dissertation, in which she examines how primary school teachers can utilize educational technology to facilitate interaction between their students in an online setting.

Raúl A. Barba-Martín has a PhD in Education. Lecturer in the University of León, Spain.

Daniel Bores-García has a PhD in Education. Physical Education teacher. Part-time lecturer. Research Group of Humanities and Qualitative Research in Health Science of Universidad Rey Juan Carlos (Spain).

Maren Breier is an MSc student on the Applied Neuroscience programme at the Institute of Psychiatry, Psychology & Neuroscience, King's College London, and a member of the co-production (co-pro) panel of online students that was set up to engage with the online programme team in focused discussions around new education initiatives.

Andrea Brosnan is a primary school teacher working in a mainstream primary school in County Kerry. She graduated with a First Class Honours from Mary Immaculate College in 2008 and was awarded the Mary Immaculate College Medal, the Carlisle and Blake Award and the INTO's Vere Foster Medal. These awards were in recognition of attaining first place in the Bachelor of Education Degree, first place in Education, Theory and Practice and first place in Teaching Practice and Curriculum Education. Immersing herself in classroom-based teaching and learning since 2008, Andrea has accrued significant experience teaching the middle and senior classes and in SET settings. Returning to academia in Mary Immaculate College was always a high priority for Andrea and she is currently undertaking a Masters of Education in Literacy Education. Her current phenomenological research explores the perspective of SENCOs on supporting the literacy development of children with dyslexia in Irish primary schools.

Emma Byrne graduated from the B.Ed degree programme in Mary Immaculate college in 2017. Currently working as a primary school teacher in Waterford, Ireland. In Year 2 of the Literacy M Ed in Mary Immaculate College.

Erin Byrne completed the Bachelor of Education with Mary Immaculate College in 2017 and have been teaching in a primary school setting for the past five years. For the past two years Bryne has been undertaking the Masters of Education in Literacy Education with Mary Immaculate College and is cur-

rently completing a dissertation investigating gender representation patterns in classroom libraries and how these libraries could be better put to use in an inquiry-based, critical literacy classroom environment.

Rebecca Curtin is a primary teacher with over a decade of teaching experience. Rebecca specialises in literacy education.

Roisin Donnelly is Head of School of Management, People and Organisations in the Faculty of Business, Technological University Dublin.

Gustavo González-Calvo has a PhD in Education. Lecturer in the University of Valladolid, Spain.

Victoria Grace-Bland is the Student-Staff Partnerships Manager at the University of Reading, UK.

Minna Hämäläinen, D.Sc. (Econ. & Bus.Adm.), is working as researcher at Lappeenranta- Lahti University of Technology, in Finland. Her research interests include entrepreneurship education in schools, especially managing entrepreneurship education in different school levels. In addition, she is a Teacher at Entrepreneurship Master´s program and Industrial Engineering Bachelor's program in LUT University. She has been building the Entrepreneurship Education Evaluation Tool for Universities of Applied Sciences and the MTEE – Measurement Tool for Entrepreneurship Education. Been involved at developing regional Entrepreneurship education strategies in Finland. Internationally and nationally active orientation in fostering teachers' and students' entrepreneurial skills, knowledge and attitudes.

Anja Harrison is a lecturer in Psychology (Education) at the Institute of Psychiatry, Psychology & Neuroscience, King's College London. Her background and research training are at the interface of psychology and biology. She gained a PhD in Psychoneuroimmunology from The Institute of Neuroimmunology and Multiple Sclerosis at the University Medical Centre Hamburg-Eppendorf, followed by a post-doctoral research position at the Health Psychology department at King's College London and a lectureship in Health Psychology at the University of Central Lancashire. She returned to work as a lecturer at King's College London in 2021. Anja is passionate about empowering students to evaluate their own progress, giving them confidence in their ability, and identifying and addressing needs for support. Her main career interest is to leverage her passion for interdisciplinary research, teaching and mentoring to attract, excite and train the next generation of scientists and professionals. She is a Fellow of the Higher Education Academy and CEO of The Collaborative Library.

Alexandra Hayward is an Academic Developer with a special interest in collaborative learning and assessment.

David Hortigüela-Alcalá is the Director of the Department of Specific Didactics and the Area of Didactics of Corporal Expression. Specialist in didactics of Physical Education, formative and shared assessment, and active and participatory methodologies. Author of more than 300 educational publications. bit.ly/dhortigüelacv bit.ly/inves_UBU

Colin Hughes is Head of the Graduate Business School in the Faculty of Business, Technological University Dublin.

Allán Laville teaches Clinical Psychology at the School of Psychology and Clinical Language Sciences (SPCLS), University of Reading. Allán has been with the University since April 2011, when he joined as a PWP Clinical Educator in the Charlie Waller Institute (CWI). By December 2013, he had progressed into the role of Senior PWP Clinical Educator working across CWI and SPCLS. He was appointed a Lecturer in Clinical Psychology in July 2019, promoted to Associate Professor of Clinical Psychology in August 2021, and most recently, promoted to Professor of Equity in Psychology. Allán is a National Teaching Fellow and Senior Fellow of the Higher Education Academy and an Associate Fellow of the British Psychological Society (BPS). He is Co-Chair of the Equity, Equality, Diversity and Inclusion Committee and member of the Scientific Programme Advisory Group for the British Association for Behavioural and Cognitive Psychotherapies. He is also a member of the BPS Equality, Diversity and Inclusion Strategic Board. In 2019, Allán won the Reading Students' Union Award for Diverse and Inclusive Teaching Excellence and in 2020, he was awarded a University Teaching Fellowship. Allán was a finalist in the BPS and Oxford University Press Higher Education Psychology Teacher of the Year 2020 and 2022 competitions.

Milena Marinkova is a Lecturer in EAP at the Language Centre of the University of Leeds, teaching on in-sessional and pre-sessional content-based EAP modules (at the School of Education), as well as credit-bearing modules on language, literature and place. An active member of the Embracing Linguistic Diversity and Pedagogical Linguistics satellites of Language@Leeds, Milena is interested in investigating the linguistic and cultural diversity in EAP / language teaching, the role of affect and creativity in EAP, and inclusive teaching and assessment practices across HE sector.

Alexandru Mihaila has a PhD. Lecturer MIHĂILĂ Alexandru Robert is a tenured professor of the Department for Teachers Training from the Bucharest University of Economics Studies. He is a graduate of the Bucharest University of Economic Studies, with a double specialization in Management and Economic Informatics and has a PhD degree in Cybernetics and Statistics. He carries out his teaching activity within the Department for Teacher Training, contributing both to the training of future teachers and to the improvement of teachers in pre-university education. He has a rich activity in educational projects, especially in adult training and student counseling activities. He has written and published specialized works, articles and studies in the educational and social field. He coordinated and carried out, together with the DPPD team, studies on: employers' requirements for graduates' competencies, the insertion of graduates of economic higher education in the labor market, students' expectations in relation.

Emma O'Brien is an Lecturer in Taught Postgraduate Programmes in Education in Mary Immaculate College and is the programme co-ordinator for the MEd in Digital Leadership in Education. Emma has worked in Digital Education since 2000 and has over 17 years experience teaching in the University of Limerick, Limerick Institute of Technology, NUIG and UCC, through blended, online and face-to-face mediums. She has taught on undergraduate, postgraduate and professional programmes and has developed innovative curricula to support the needs of a variety of learners, particularly those requiring flexible models of learning to support them to return to education. Emma has supervised 125 final year students and 24 masters students. She has a PhD in Digital learning (2005, UL) and a Masters in Computing in Education (2003, LIT). Emma is an accomplished researcher and has been Principle Investigator (PI) on several large scale European research projects. Her total funding income as PI amounts to over €1 million in the areas of digital learning, workplace learning, problem and inquiry based learning with over

50 publications in these fields. Emma was previously co-chair of FACILITATE - the national inquiry based learning network and is currently external examiner for work based learning at the University of Northampton.

T.J. Ó Ceallaigh is an Associate Professor at the School of Education, University College Cork, Ireland. His main research interests focus on the pedagogy required for the successful integration of language and content instruction and assessment, with particular reference to language immersion and bilingual contexts. T.J. is co-chair of Teacher Education and Digital Technology Research and Development Community for the Association for Teacher Educators in Europe (ATEE) and is a Board Director of Integrating Language and Content in Higher Education (ICLHE), an international association of scholars who are interested in and actively contribute to the multi-disciplinary field of content and language integration in higher education.

Nick Pilcher is a lecturer in The Business School at Edinburgh Napier University (UK). He lectures and helps students with subject assignments. He researches in a range of areas. He has published and contributed to articles in journals including Teaching in Higher Education and Maritime Policy and Management.

Harriet Power is an MSc student on the Psychology and Neuroscience of Mental Health programme at the Institute of Psychiatry, Psychology & Neuroscience, King's College London, and a member of the co-production (co-pro) panel of online students that was set up to engage with the online programme team in focused discussions around new education initiatives.

Anu Raappana works as a Junior researcher at LUT University in Finland. Her research focuses on entrepreneurship and entrepreneurship education. The current study examines young people's entrepreneurial skills as well as adolescents' perceptions of their own entrepreneurial potential.

Elsa Ribeiro-Silva holds a PhD in Sport Sciences and Physical Education from the University of Coimbra. She is also the Coordinator of the Masters Degree of Teaching Physical Education and the Branch of Physical Education of the PhD in Sport Sciences.

Kendall Richards lectures in the School of Computing at Edinburgh Napier University (UK). His research interests are in education as social justice and language. He has published and contributed to articles published in journals including International Journal of Qualitative Studies in Education, Teaching in Higher Education, and Power and Education.

Elena-Ramona Richiteanu-Nastase is a university lecturer at the Bucharest University of Economic Studies, Teachers Training Department, has a PhD in Educational Sciences, over 15 years of experience in education / training / research/ counseling: national and international educational projects and programs or collaborations within NGOs (Romanian Center for Economic Education, Education 2000+), individual counseling and group counseling (students and their families, teachers, unemployed), 15 years of teaching experience at the Teachers Training Department (course holder for subjects such as Counseling and Guidance, Curriculum Theory and Methodology, Training Theory and Methodology, Evaluation Theory and Methodology,) and 6 years as a trainer of the National Institute of Magistracy

(Personal Development and Communication, head of department). She has a rich publishing activity: book as sole author (Counselling and guidance for career and life Theoretical and methodological foundations) or in collaboration (over 10), articles and studies (over 30).

Joy Robbins is an Associate Professor in English for Academic Purposes (EAP) and a Senior Fellow of the Higher Education Academy/ Advance HE. She teaches at the Language Centre at the University of Leeds, UK and has interests in digitally enhanced learning, assessment, and the teaching and learning of writing.

Martin Sands is a Lecturer in Medical Education at King's College London. He has a postgraduate MSc in Physiotherapy and Education, including a PGCert in Practice Based Education and is a Fellow of the Higher Education Academy. He worked as a clinical physiotherapist in UK health systems for over 15 years before developing his role as an educationalist in UK Higher Education, including over 8 years experience of curriculum design, development and implementation. His research interests include reflection and reflective practices in healthcare education.

Murat Tekin, MD., is an Assistant Professor of Medical Education at Çanakkale Onsekiz Mart University. He teaches Family Medicine courses at the faculty and is responsible for the skill training of the faculty. His areas of interest are medical education, teaching methods and techniques.

Brenda Williams is a Reader in Neuroscience Education at the Institute of Psychiatry, Psychology & Neuroscience (IoPPN), King's College London and a senior fellow of the HEA. She works to enhance the student experience through evaluating and improving assessment and feedback practices and promoting skills development through her work as Assessment and Feedback Lead and Faculty Assessment Board Chair. Being a firm believer in diversity and inclusion, she led the development two successful distance learning MSc programmes at the IoPPN: Psychology and Neuroscience of Mental Health, and Applied Neuroscience. With a research background in the neurosciences, she is passionate about working with colleagues and students to promote and integrate innovative neuroscience education.

Yue Yue, PhD, FHEA, is a Lecturer in Psychology at the University of Reading. She teaches Foundation and Undergraduate Psychology Courses, including skill-based modules (academic skills, research methods) and fundamental psychology topics. She is a Cognitive Psychologist, and her current interest is to explore collaborative learning. She is a member of the University of Reading Decolonising the Curriculum working group.

Index