



Anne Piezunka

Ist eine gute Schule eine inklusive Schule?

Entwicklung von Messinstrumenten durch Schulinspektionen

Piezunka

Ist eine gute Schule eine inklusive Schule?

Interdisziplinäre Beiträge zur Inklusionsforschung

herausgegeben von Marina Egger, Julia Frohn,
Vera Moser und Detlef Pech

Anne Piezunka

Ist eine gute Schule eine inklusive Schule?

Entwicklung von Messinstrumenten
durch Schulinspektionen

Verlag Julius Klinkhardt
Bad Heilbrunn • 2020

k

Meiner Familie
Hille, Hartmut, Henning & Christian

*Die Publikation wurde durch den Leibniz-Publikationsfonds für Open-Access-Monografien gefördert.
Dieser Titel steht als Klinkhardt Open Access im Fachportal Pädagogik zum Download zur Verfügung.*

*Die vorliegende Arbeit wurde 2019 mit dem Wissenschaftspreis
der Sektion Sonderpädagogik der DGfE ausgezeichnet.*

Die vorliegende Arbeit wurde von dem Institut Erziehungswissenschaften der Universität Hildesheim unter dem gleichnamigen Titel als Dissertation angenommen.

Gutachter: Prof. Dr. Vera Moser, Prof. Dr. Melanie Fabel-Lamla.

Tag der Disputation: 9.05.2019.

Dieser Titel wurde in das Programm des Verlages mittels eines Peer-Review-Verfahrens aufgenommen.
Für weitere Informationen siehe www.klinkhardt.de.

Bibliografische Information der Deutschen Nationalbibliothek

Die Deutsche Nationalbibliothek verzeichnet diese Publikation
in der Deutschen Nationalbibliografie; detaillierte bibliografische Daten
sind im Internet abrufbar über <http://dnb.d-nb.de>.

2020.kg. © by Julius Klinkhardt.

Grafik Umschlagseite 1: © kate_kalita / adobe stock.

Druck und Bindung: AZ Druck und Datentechnik, Kempten.

Printed in Germany 2020.

Gedruckt auf chlorfrei gebleichtem alterungsbeständigem Papier.



*Die Publikation (mit Ausnahme aller Fotos, Grafiken und Abbildungen) ist veröffentlicht unter der
Creative Commons-Lizenz: CC BY-NC-SA 4.0 International
<https://creativecommons.org/licenses/by-nc-sa/4.0/>*

ISBN 978-3-7815-5799-4 Digital

doi.org/10.35468/5799

ISBN 978-3-7815-2360-9 Print

Zusammenfassung

Die Umsetzung von schulischer Inklusion findet in Deutschland auf verschiedenen Ebenen des Bildungssystems statt, z.B. in der Gestaltung der Schulstruktur. In diesem Zusammenhang stellt sich die Frage, wie sich die Umsetzung von Inklusion abbilden lässt. In Abhängigkeit davon, wie Inklusion als politische Vorgabe gedeutet wird, kann es beispielsweise variieren, welche empirischen Daten hierfür erhoben werden. Daran anknüpfend geht es im Rahmen der Arbeit um die Fragen, inwiefern Inklusion in den Messinstrumenten der Schulinspektion berücksichtigt wird und wie diese entstanden sind. Schulinspektionen evaluieren in den einzelnen Bundesländern mithilfe von Messinstrumenten, z.B. Schülerfragebögen und Unterrichtsbeobachtungsbögen, die Qualität von Einzelschulen. Dabei unterscheiden sich die Bundesländer darin, welche Instrumente sie anwenden. Durch die Analyse der Messinstrumente und mithilfe von Experteninterviews wird herausgearbeitet, ob und welche Daten in Bezug auf die Umsetzung von Inklusion erhoben werden. Des Weiteren wird deutlich, wie die Messinstrumente der Schulinspektionen entstanden sind und welche Rahmenbedingungen prägend wirkten. Die Auseinandersetzung mit dem Entwicklungsprozess ermöglicht es zu zeigen, warum sich die Messinstrumente zwischen den Bundesländern unterscheiden. Für die Bearbeitung der Fragestellung wird auf organisations- und bewertungssoziologische Arbeiten (z.B. Sensemaking Ansatz) zurückgegriffen sowie auf zentrale Konzepte der Inklusionsforschung und den Educational Governance-Ansatz.

Abstract

Due to the UN Convention on the Rights of Persons with disabilities, the Federal States of Germany have to implement inclusive education. However, there is no shared understanding how to define inclusive education in Germany. Therefore, different understandings on how to implement inclusive education within the schools exist. Comparing different Federal States of Germany, Anne Piezunka examines how school inspectorates define inclusive education. School inspectorates are part of the educational administration. They evaluate the quality of schools and are a key means of accountability. Evaluating the quality of schools, they analyze documents, conduct interviews and observe classes. Thereby, they apply several metrics. For example: "The class starts on time". Every federal state has developed its own metrics. Piezunka analyzes which metrics school inspectorates apply to evaluate the implementation of inclusive education on the school level. Interviewing the developers of metrics, she also studies the development of these metrics and seeks to explain differences between the federal states. The study brings together sensemaking theory, sociology of evaluation and research on inclusive education.

Zusammenfassung orientiert an leichter Sprache

Was ist eine gute Schule? Auf diese Frage gibt es viele Antworten. Zum Beispiel:

- Wenn Lehrer_innen nett zu den Kindern sind.
- Wenn die Klassenzimmer sauber sind.
- Wenn Kinder an der Schule mitbestimmen können.
- Wenn es Computer an der Schule gibt.

Es gibt noch viele weitere Antworten. In meiner Arbeit beschäftige ich mich mit Schulinspektor_innen. Was machen sie? Sie besuchen einzelne Schulen und machen folgendes:

- Sie sitzen mit den Kindern im Klassenzimmer und beobachten den Unterricht.
- Sie unterhalten sich mit der Schulleitung und mit Lehrer_innen.
- Sie unterhalten sich mit den Kindern.

Nach ihrem Besuch schreiben die Schulinspektor_innen einen Bericht. In dem Bericht steht, was die Schule gut macht. In dem Bericht steht auch, was die Schule schlecht macht. Warum machen sie das? Sie wollen der Schule helfen. Durch den Bericht weiß die Schule, was sie verbessern muss.

Ich will wissen: Was ist für Schulinspektor_innen eine gute Schule? Wenn sie eine Schule besuchen, können sie nämlich nicht auf alles achten. Sie haben nicht genug Zeit. Deswegen müssen sie Entscheidungen treffen. Zum Beispiel: Ist es ihnen wichtig, dass der Schulflur sauber ist? Oder finden sie es wichtig, dass es in jedem Klassenzimmer Computer gibt?

In den letzten Jahren habe ich viel gelesen: Es gibt viele Texte von den Schulinspektor_innen. In den Texten beantworten die Schulinspektor_innen die Frage: Was ist für uns eine gute Schule? Ich habe die Texte gelesen und ich habe mit Schulinspektor_innen gesprochen. Ich habe die Schulinspektor_innen gefragt: Findet ihr das wichtig,

- ob die Lehrer_innen nett zu allen Kindern sind.
- dass die Kinder mitbestimmen können.
- ob alle Kinder beim Lernen die richtige Hilfe erhalten.

Auf meine Fragen habe ich unterschiedliche Antworten bekommen. Dann hatte ich eine weitere Frage: Warum antworten die Schulinspektor_innen so unterschiedlich? Das habe ich auch die Schulinspektor_innen gefragt. Dafür gibt es verschiedene Gründe:

- Schulinspektor_innen leben an unterschiedlichen Orten in Deutschland.
- Die Schulinspektor_innen haben in ihrem Leben unterschiedliche Erfahrungen gemacht. Zum Beispiel: Sie haben an unterschiedlichen Schulen gearbeitet.
- Die Schulinspektor_innen haben mit Kolleg_innen, Bekannten und ihrem Chef darüber gesprochen, was eine gute Schule ist. Dabei haben sie unterschiedliche Antworten erhalten.

Deswegen haben die Schulinspektor_innen unterschiedliche Antworten gegeben. Im folgenden Text habe ich dies ausführlich erklärt.

Inhaltsverzeichnis

1	Einleitung	13
1.1	Forschungsfragen	15
1.2	Relevanz der Arbeit	15
1.2.1	Analyse von Messinstrumenten	15
1.2.2	Implementation von Inklusion	17
1.3	Verortung der Untersuchung	17
1.4	Aufbau der Arbeit	19
2	Schulische Inklusion – eine Begriffsklärung	21
2.1	Verständnisse von schulischer Inklusion	21
2.2	Menschenrechtliches Verständnis	24
2.2.1	Menschenrechtliche Perspektive	24
2.2.2	Fokus auf die Einzelschulebene	25
2.2.3	Teilhabe und Nichtdiskriminierung in Bezug auf Schule	26
2.2.4	Inklusive Perspektive auf Menschenrechte	30
2.2.5	Abgrenzung zu anderen Verständnissen	35
2.3	Forschungsstand: Umsetzung von Inklusion in Deutschland	36
2.3.1	Individuell bestmögliche Qualifizierung	37
2.3.2	Erfahren von intersubjektiver Anerkennung	38
2.3.3	Politische Partizipation	39
2.3.4	Zwischenfazit	40
3	Schulinspektionen in Deutschland	43
3.1	Beschreibung eines Inspektionsverfahrens	43
3.2	Einführung von Schulinspektionen	44
3.3	Schulinspektionen im Bildungssystem	47
3.3.1	Educational Governance-Ansatz	47
3.3.2	Schulinspektion als Teil der öffentlichen Verwaltung	48
3.3.3	Funktionen der Schulinspektion	49
3.4	Zwischenfazit	53
3.5	Forschungsstand zu Schulinspektionen	53
3.5.1	Schulinspektionen als ein Akteur im Bildungssystem	54
3.5.2	Wirkung von Schulinspektionen	54
3.5.3	Analyse der Messinstrumente	56
3.5.4	Entwicklung der Messinstrumente	57
4	Messinstrumente aus bewertungssoziologischer Perspektive	59
4.1	Merkmale von Indikatoren	60
4.1.1	Indikatoren als Klassifizierungen	60
4.1.2	Indikatoren als eine Form der Bewertung	61
4.1.3	Kommensurierung	62
4.1.4	Vergleiche von Entitäten	63

4.2	Entwicklung von Bewertungsinstrumenten	65
4.2.1	Forschungsstand	65
4.2.2	Bewertungskonstellationen	66
4.2.3	Institutionalisierung von Indikatoren	67
4.3	Exkurs: Qualität von Schule	71
4.4	Zwischenfazit	73
5	Sensemaking-Ansatz	75
5.1	Sensemaking-Ansatz nach Weick	75
5.1.1	Kennzeichen von Sensemaking nach Weick	77
5.1.2	Sensemaking in Organisationen	79
5.1.3	Sensemaking und Macht	80
5.2	Implementation von politischen Vorgaben	81
5.2.1	Abgrenzung zu anderen Ansätzen	81
5.2.2	„Distributed-Cognition“-Ansatz	83
5.2.3	Zwischenfazit	86
6	Empirisches Design	87
6.1	Analyse der Bewertungsinstrumente	87
6.1.1	Datenkollektion	87
6.1.2	Relevante Dokumente	89
6.1.3	Qualitative Inhaltsanalyse der Bewertungsinstrumente	93
6.2	Analyse der Experteninterviews	97
6.2.1	Entwicklung des Leitfadens und theoretische Orientierung	98
6.2.2	Kontaktaufnahme zu den Interviewpartnern	99
6.2.3	Durchführung der Interviews	99
6.2.4	Reflexion der Interviews	100
6.2.5	Qualitative Inhaltsanalyse der Experteninterviews	103
6.2.6	Grenzen der Datenanalyse	105
6.3	Forschungsethik und Datenschutz	107
6.3.1	Informierte Einwilligung	107
6.3.2	Datenschutz	107
7	Deutungen von Inklusion	109
7.1	Berücksichtigung von Konzepten in Bezug auf Inklusion	110
7.1.1	Ziel von Inklusion: Individuell bestmögliche Qualifizierung	110
7.1.2	Ziel von Inklusion: Politische Partizipation aller	115
7.1.3	Ziel von Inklusion: Erfahren von Anerkennung	117
7.1.4	Nennung von Differenzlinien	119
7.1.5	Berücksichtigung von Inklusion als normative Forderung	125
7.2	Verwendung des Inklusionsbegriffs	126
7.2.1	Verständnis von Inklusion	126
7.2.2	Instrumente in Bezug auf Inklusion	130
7.2.3	Zwischenfazit: Verständnisse von Inklusion in den Instrumenten	134

7.3	Inklusionsdeutungen der Instrumenteentwickler_innen	136
7.3.1	Typ I: Inklusion als Anspruch auf Förderung	136
7.3.2	Typ II: Inklusion als Herausforderung für Schulen	139
7.3.3	Typ III: Inklusion als Teilaspekt von guter Schule	140
7.4	Deutungen von Inklusion im Vergleich	142
7.5	Exkurs: Konsensuales Verständnis von Schulqualität	142
8	Entwicklung von Messinstrumenten zu Inklusion	145
8.1	Prägende Rahmenbedingungen beim Sensemaking	147
8.1.1	Persönliche Attribute der Entwickler_innen	147
8.1.2	Rechtliche Vorgaben	152
8.1.3	Situativer Kontext	154
8.1.4	Beispiel: Bewerten und Beschreiben in der Datenauswertung	174
8.2	Relevante Impulse beim Sensemaking	177
8.2.1	Sensemaking von Teilaspekten von Inklusion	178
8.2.2	Sensemaking zu Lernstandserhebungen	179
8.3	Sensemaking von einzelnen Fällen	180
8.3.1	Frau Elsner als Vertreterin von Typ I	181
8.3.2	Herr Pehla als Vertreter von Typ II	184
8.3.3	Herr Plate und Frau Steigleder als Vertreter von Typ III	185
8.4	Kollektive Aushandlungsprozesse	191
8.5	Einzelfallübergreifende Handlungsmuster	192
9	Diskussion und Ausblick	195
9.1	Berücksichtigung von Inklusion	195
9.1.1	Konzepte in Bezug auf Inklusion	196
9.1.2	Verständnisse von Inklusion in den Instrumenten	202
9.1.3	Schlussfolgerungen aus dem zweistufigen Analyseverfahren	206
9.1.4	Deutungen von Seiten der Instrumenteentwickler_innen	207
9.2	Entwicklung von Indikatoren zu Inklusion	209
9.2.1	Kurzzusammenfassung der Ergebnisse aus der Analyse	209
9.2.2	Gründe für unterschiedliche Deutungen	210
9.2.3	Implikationen der jeweiligen Rahmenbedingungen	212
9.2.4	Bewertungssoziologische Überlegungen	217
10	Fazit	223
	Literatur	227
	Abbildungsverzeichnis	241
	Tabellenverzeichnis	241
	Akronyme	243

A Leitfaden	245
A.1 Teil 1: Allgemeine Fragen	246
A.2 Teil 2: Fragen zu Instrumenten	248
A.3 Teil 3: Bundesländerspezifische Fragen	250
B Transkriptionsregeln	251
C Danksagungen	253

1 Einleitung

„Evaluation is a cultural process that—broadly defined—concerns the negotiation, definition and stabilization of value in social life“ (Lamont et al., 2014)¹.

Es variiert nicht nur zwischen Staaten und Regionen, sondern auch zwischen Einzelpersonen, was unter einer „guten“ Schule verstanden wird. Dabei unterscheiden sich die einzelnen Vorstellungen von Schulqualität beispielsweise darin, wie die gesellschaftlichen Funktionen von Schule nach Fend (2006) gedeutet werden und welche Relevanz ihnen zugeschrieben wird. Obwohl verschiedene Verständnisse von Schulqualität existieren (vgl. Ditton 2000, S. 74), entsteht im öffentlichen Diskurs häufig der Eindruck, dass Konsens darüber herrscht, was eine gute Schule sei (vgl. Heid 2007, S. 55). So wird beispielsweise im Fall von datengestützten Evaluationen von Bildungseinrichtungen häufig argumentiert, dass hierbei auf ein allgemeingültiges Verständnis von guter Schule zurückgegriffen werden würde (nähere Beschreibung siehe Heinrich 2010, S. 47). Demnach werden die Daten, die mithilfe der Messinstrumente erhoben werden, als objektive Grundlage angesehen (Sellar 2014), um evidenzbasierte Entscheidungen treffen zu können.

Jedoch repräsentieren die Instrumente, die zur Messung von Qualität genutzt werden, stets bestimmte Vorstellungen von guter Schule und vernachlässigen andere. Beispielsweise kann der Fokus auf den erworbenen Kompetenzen von Schüler_innen liegen oder es steht die Frage im Mittelpunkt, ob im Sinne von Inklusion die Teilhabe und Nichtdiskriminierung aller Schüler_innen an der Einzelschule gewährleistet ist. Bei der Entwicklung von Messinstrumenten können demnach unterschiedliche Schwerpunkte gesetzt werden bzw. basieren die einzelnen Verständnisse von guter Schule zum Teil auf unterschiedlichen Wertvorstellungen. Daher lässt sich schlussfolgern, dass Messinstrumente nicht neutral sind, sondern stets eine normative Komponente beinhalten (vgl. Biesta 2015, S. 316). Demnach bedarf es normativer Setzungen, wenn Messinstrumente entwickelt werden und diese sind nicht per se naturgegeben (vgl. Bowker und Star 2000, S. 5), sondern sozial konstruiert.

Diese normative Komponente gerät jedoch häufig in Vergessenheit, wenn die Instrumente bereits länger in Gebrauch sind und von vielen genutzt werden: „As they [metrics, ap] simplify and reduce complex educational processes to numerical ratings and rankings, however, these metrics can also narrow debate and the kind of education one think is possible“ (Anagnostopoulos et al. 2016, S. 343).

Es ist diese normative Komponente von Messinstrumenten, die im Rahmen dieser Arbeit näher betrachtet wird. So wird untersucht, wie ein Teilaspekt von Schulqualität, Inklusion, in den Messinstrumenten der Schulinspektionen berücksichtigt wird. Der Fokus auf Inklusion bietet sich an, weil Deutschland sich unter anderem durch die UN-Behindertenrechtskonvention (BRK) zur Umsetzung von Inklusion verpflichtet hat. Auch wenn Inklusion in der Öffentlichkeit intensiv diskutiert wird, gibt es kaum Wissen darüber, ob und wie die politische Vorgabe auf den einzelnen Ebenen des Bildungssystems implementiert wurde. Dabei wird die Ebene der Bildungsadministration im Rahmen der Arbeit am Beispiel der Schulinspektionen² näher

¹ Lamont et al. verweisen hierbei auf die theoretischen Überlegungen von Beckert und Musselin (2013).

² Im Folgenden werden externe Evaluationen von Schulen als Schulinspektionen bezeichnet. Die Bundesländer verwenden hierfür unterschiedliche Bezeichnungen, z. B. Fokusevaluation, Qualitätsanalyse, Externe Evaluation, usw. (vgl. Sowada und Dederling 2016).

betrachtet. So sind Schulinspektionen in den einzelnen Bundesländern dafür zuständig, mithilfe von empirischen Erhebungsverfahren die Qualität von Einzelschulen zu evaluieren. Hierbei greifen sie beispielsweise auf Unterrichtsbeobachtungsbögen oder Onlinefragebögen zurück, um Daten zu erheben. Dabei sind Schulinspektionen nur eine von vielen Maßnahmen, die seit der Jahrtausendwende eingeführt wurden, um im Rahmen von „neuer Steuerung“ mehr Daten über Schulen zu sammeln, um diese für bildungspolitische Entscheidungen nutzen zu können (vgl. Bellmann, Duzevic et al. 2016).

In Bezug auf die bundesländerspezifischen Messinstrumente der Schulinspektionen und deren Inhalte zeigt sich, dass es hierzu bislang kaum empirische Untersuchungen gibt (Ausnahme: De-genhardt 2008). Dies ist insofern überraschend, als sich aus den Messinstrumenten ablesen lässt, welche Anforderungen die Bildungsadministration in den jeweiligen Bundesländern an die Einzelschule formuliert hat. Die Arbeit geht noch einen Schritt darüber hinaus und zeigt am Beispiel „Inklusion“, welche Rahmenbedingungen die Entwicklung der Messinstrumente prägen. Dadurch soll abgebildet werden, wieso sich bestimmte Wertvorstellungen in den Messinstrumenten durchsetzen und andere nicht. Hierbei wird auch untersucht, inwiefern sich die Schulinspektionen in den einzelnen Bundesländern in ihren Deutungen von Inklusion sowie in der Entwicklung der Instrumente zu Inklusion unterscheiden.

Im Hinblick auf das Thema der Arbeit ist zu berücksichtigen, dass im deutschsprachigen Raum kein Konsens darüber existiert, wie Inklusion im schulischen Kontext definiert wird (vgl. Moser 2012b; Grosche 2015). Des Weiteren gibt es im öffentlichen Diskurs auch unterschiedliche Positionen zu der Frage, ob die Umsetzung von Inklusion erstrebenswert bzw. ein Teilaspekt von Schulqualität sei (z. B. Olbrisch 2017). Um Aussagen darüber treffen zu können, ob Inklusion in den Messinstrumenten der Schulinspektion verankert ist, ist es notwendig, in Bezug auf die beiden oben aufgeführten Aspekte bestimmte Setzungen vorzunehmen: Inklusion wird daher im Rahmen dieser Arbeit als ein Teilaspekt von Schulqualität verstanden und Inklusion wird aus menschenrechtlicher Perspektive als das Ermöglichen von Teilhabe und die Nichtdiskriminierung aller Schüler_innen definiert (nähere Ausführung → Kapitel 2).

Zugleich wird durch ein zweistufiges Verfahren der Aspekt, dass es kein einheitliches Verständnis von Inklusion gibt, im empirischen Design aufgegriffen: So wird näher betrachtet, wie von Seiten der Schulinspektionen Inklusion als Begriff gedeutet wird bzw. wie sie sich dazu positionieren. Diese zusätzliche Perspektive birgt diverse Potenziale: Zum einen kann beantwortet werden, inwiefern Inklusion in den einzelnen Bundesländern auf der Ebene der Bildungsadministration unterschiedlich definiert wird und inwiefern daher auch voneinander abweichende Anforderungen an die jeweiligen Schulen formuliert werden. Zum anderen kann untersucht werden, ob das oben beschriebene Verständnis von guter Schule im Sinne von Teilhabe und Nichtdiskriminierung von Seiten der Schulinspektionen mit dem Begriff „Inklusion“ verknüpft wird bzw. ob jene Wertvorstellung unabhängig vom Inklusionsbegriff sichtbar wird. Demnach leistet diese Arbeit einen Beitrag zu der Frage, wie Messinstrumente entstehen und inwiefern Inklusion als Verständnis von guter Schule in den Instrumenten berücksichtigt wird. Zugleich ermöglicht die Arbeit auch Aussagen zur Implementation von Inklusion als politische Vorgabe in den einzelnen Bundesländern.

1.1 Forschungsfragen

Die folgenden Forschungsfragen stehen daher im Zentrum der Analyse:

1. Inwiefern wird Inklusion in den Messinstrumenten der Schulinspektionen berücksichtigt?
 - Inwiefern wird in den Instrumenten der Schulinspektionen ein Verständnis von guter Schule sichtbar, wonach im schulischen Alltag die Teilhabe und Nichtdiskriminierung aller (-Inklusion) angestrebt wird?
 - Welche Verständnisse von Inklusion spiegeln sich in denjenigen Instrumenten wider, die von Seiten der Schulinspektionen als Instrumente für die Messung von Inklusion gekennzeichnet sind bzw. wie wird Inklusion von den einzelnen Instrumenteentwickler_innen gedeutet?
 - Inwiefern unterscheiden sich die bundesländerspezifischen Schulinspektionen in ihren Deutungen von Inklusion bzw. in den dazugehörigen Messinstrumenten?
2. Wie sind die Instrumente, die dazu dienen, die Umsetzung von Inklusion zu messen, zustande gekommen?
 - Welche Rahmenbedingungen haben die Deutung von Inklusion sowie die Entwicklung von Instrumenten geprägt?
 - Welche Handlungsmuster werden bei der Entwicklung von Instrumenten zur Messung von Qualität sichtbar?

1.2 Relevanz der Arbeit

Im Folgenden wird die Relevanz, sich mit den Inhalten und der Entwicklung von Instrumenten zur Messung von Schulqualität sowie der Implementation von Inklusion auseinanderzusetzen, dargestellt.

1.2.1 Analyse von Messinstrumenten

Unabhängig davon, ob Messinstrumente zur Deskription oder zur Bewertung von Objekten genutzt werden, leisten sie einen Beitrag dazu, wie Objekte wahrgenommen werden. Sie konstruieren die Wahrnehmung der Wirklichkeit, indem durch ihre Verwendung Klassifizierungen von Objekten vorgenommen werden (vgl. Durkheim und Mauss 1987). Diese können in Bezug auf Ungleichheiten relevant sein, wenn eine Hierarchisierung vorgenommen wird oder wenn bestimmte Aspekte nicht Teil der Klassifikation sind. Beispielsweise kritisiert Supik (2017), dass bei der Erhebung von Gender meistens weiterhin auf eine dichotome Unterscheidung zwischen Mann und Frau zurückgegriffen wird. Somit werden durch solche Messformen heteronormative Strukturen reproduziert und bestimmte Personengruppen, z. B. Intersexuelle, werden in statistischen Erhebungen nicht sichtbar. Darüber hinaus kann das Fehlen von bestimmten Klassifikationen implizieren, dass systematische Formen der Diskriminierung überhaupt nicht sichtbar werden. Beispielsweise gibt es in der Literatur erste Hinweise, dass Menschen mit hohem Körpergewicht im schulischen Kontext diskriminiert werden (z. B. Helbig und Jähnen 2013). In den meisten empirischen Studien fehlt jedoch eine Erhebung dieses Merkmals. Messinstrumente treffen daher Setzungen, die sich darauf auswirken, wie Wirklichkeit wahrgenommen wird.

Im Rahmen dieser Arbeit geht es um Messinstrumente, die nicht nur zur Deskription von Objekten, sondern explizit zur Evaluation genutzt werden, d. h. es geht nicht nur darum, eine

Form von Wirklichkeit abzubilden, sondern die Daten dienen darüber hinaus dazu, Aussagen über die Qualität eines Objekts machen zu können. Die Entscheidung für bestimmte Vorstellungen von Schulqualität und die dazugehörigen Messinstrumente impliziert eine Entscheidung gegen andere. So können im Rahmen einer Evaluation aus pragmatischen Gründen, z. B. finanzielle und zeitliche Ressourcen, nicht alle Qualitätsvorstellungen abgedeckt werden bzw. gibt es auch Qualitätsvorstellungen, die nicht miteinander kompatibel sind. Die Abgrenzung von anderen Wertüberzeugungen kann demnach eine bewusst getroffene Entscheidung sein, um sich gegen bestimmte Qualitätsvorstellungen zu positionieren. Jene Entscheidung kann auch unbewusst getroffen werden, indem bei der Messung von Qualität andere Schwerpunkte gesetzt werden und dies indirekt zur Folge hat, dass bestimmte Qualitätsvorstellungen nicht repräsentiert sind. Die Frage danach, welche Messinstrumente verwendet werden, wirkt sich bei der Durchführung von Evaluationen darauf aus, wie ein Objekt bewertet wird. Beispielsweise zeigt sich bei Universitätsrankings, dass Veränderungen in den Auflistung der Hochschulen unter anderem auf Veränderungen bei der Auswahl der Messinstrumente zurückzuführen sind (vgl. Espeland und Sauder 2007).

In Bezug auf das Bildungssystem in Deutschland liegt es nahe zu untersuchen, inwiefern das Ermöglichen von Teilhabe und Nichtdiskriminierung aller Schüler_innen als Verständnis von guter Schule in den Messinstrumenten der Schulinspektionen berücksichtigt wird. Diskriminierung kann sich im schulischen Setting unter anderem in der Unterrichtsgestaltung durch fehlende Formen der Differenzierung, in der Interaktion zwischen Lehrkraft und Schüler_innen (vgl. Prengel 2013) oder in Selektionsprozessen (z. B. Radtke und Gomolla 2009) widerspiegeln. Eine Vielzahl von Studien belegt, dass bestimmte Gruppen von Schüler_innen im Bildungssystem diskriminiert werden, z. B. Schüler_innen mit „Migrationshintergrund“ (z. B. Gomolla und Radtke 2002; Sprietsma 2009; Diefenbach 2010; Diehl et al. 2016) oder Schüler_innen mit diagnostizierter Behinderung (z. B. Powell 2006; Pfahl 2011). Demnach besteht im Fall von Deutschland bezüglich Teilhabe und Diskriminierung im schulischen Setting Handlungsbedarf. Hierbei stellt sich die Frage, ob diese Problematik auch von Seiten der Schulinspektionen aufgegriffen wird oder ob durch die Messinstrumente andere Verständnisse von guter Schule in den Blick genommen werden. So zeigt sich beispielsweise als globaler Trend, dass das Erreichen von bestimmten Kompetenzen den Diskurs um Schulqualität dominiert (vgl. Heinrich 2010, S. 47; nähere Ausführung folgt → Kapitel 2). Diedrich kritisiert in dem Zusammenhang: „Es entsteht der Eindruck, Bildungsqualität werde auf ein ganz spezifisches Verständnis reduziert, das sich primär am Maßstab der fachlichen Leistungsfähigkeit in den Kernfächern (...) orientiert“ (Diedrich 2008, S. 14). Daran anknüpfend soll im Rahmen der Arbeit thematisiert werden, wo sich die Messinstrumente der deutschen Schulinspektionen verorten lassen, d. h. auf welche Verständnisse von Schulqualität zurückgegriffen wird.

Bislang gibt es nicht nur in Bezug auf Schulinspektionen, sondern generell kaum Forschung, die sich mit der Entwicklung von Messinstrumenten auseinandersetzt (z. B. Desrosières 2009). Dies ist insofern überraschend, da in verschiedenen Bereichen standardisierte Instrumente für Datenerhebungen genutzt werden, z. B. bei Erhebungen durch das Statistische Bundesamt, im Dienstleistungsbereich bei der Evaluation von Produkten oder bei Evaluationen durch die empirische Bildungsforschung. Die Daten, die in diesem Rahmen entstehen, dienen häufig als Grundlage, um weitreichende Entscheidungen zu treffen, z. B. Ressourcenverteilungen oder Einführung von politischen Maßnahmen. Demnach ist es nicht nur im Bildungsbereich, sondern auch in anderen gesellschaftlichen Bereichen von Relevanz zu untersuchen, wie Messinstrumente zustande gekommen sind und welche Wertüberzeugungen sich im jeweiligen Kontext

durchgesetzt haben. Ausgehend von der vermeintlichen Objektivität, die Messinstrumenten häufig zugeschrieben wird, bietet deren Analyse auch die Möglichkeit „to critically re-examine normative issues and to redefine problems that have been framed in ways so far unquestioned in public debate“ (Straßheim 2015, S. 323).

1.2.2 Implementation von Inklusion

Durch die Analyse der Messinstrumente im Rahmen dieser Arbeit lassen sich auch Aussagen über die Implementation von Inklusion als politische Vorgabe auf der Ebene der Bildungsadministration treffen. Deutschland hat sich durch die Ratifizierung der BRK zur Umsetzung von Inklusion im Bildungssystem verpflichtet. Es fehlt jedoch bislang an konkreten Vorgaben, welche Implikationen sich aus Art. 24 BRK für die Einzelschule ergeben (vgl. Wrase und Piezunka 2017) bzw. unterscheiden sich die Bundesländer auch darin, wie sie die Vorgaben der Vereinten Nationen (UN) in ihren Schulgesetzen verankert haben. Darüber hinaus gibt es aus anderen empirischen Studien bereits erste Hinweise, dass die Umsetzung von Inklusion zwischen den Bundesländern variiert (z. B. Blanck et al. 2013). So kann es sein, dass die Schulinspektionen in Bezug auf Inklusion unterschiedliche Anforderungen an Schulen formulieren, die sich in den jeweiligen Messinstrumenten widerspiegeln. Es ist auch möglich, dass die Umsetzung von Inklusion in manchen Bundesländern überhaupt nicht als Teilaspekt von guter Schule gesehen wird und Inklusion in den Messinstrumenten nicht thematisiert wird.

Es gibt zwar mehrere Studien, die sich damit auseinandersetzen, wie Inklusion auf den einzelnen Ebenen des Bildungssystems (z. B. auf der Einzelschulebene: Görtler 2015) oder auf Bundeslandebene (z. B. Nesyba 2018; Gasterstädt 2016) gedeutet wird, aber es fehlt häufig an Informationen darüber, wie es zu den jeweiligen Deutungen gekommen ist und welche Rahmenbedingungen von Relevanz waren. Darüber hinaus kann die Analyse der Messinstrumente erste Hinweise darauf geben, ob die jeweiligen Deutungen von Inklusion repräsentieren, welche Anforderungen sich durch Art. 24 BRK ergeben (vgl. Wrase 2015). Die Analyse der Rahmenbedingungen dient darüber hinaus der Praxis: Sie erhalten Informationen darüber, wie Gesetzesvorgaben auf den einzelnen Ebenen des Bildungssystems – in diesem Fall der Bildungsadministration – kommuniziert werden müssten, um stärker darauf hinzuwirken, dass diese so implementiert werden, wie es intendiert war (vgl. Coburn 2001, S. 146).

1.3 Verortung der Untersuchung

Die Studie befindet sich an der Schnittstelle verschiedenster Forschungsausrichtungen: der Bewertungssoziologie, der Implementation von politischen Vorgaben sowie der Inklusionsforschung. Die Bezüge zu diesen Forschungsausrichtungen werden im Folgenden dargestellt. Dabei wird nur kurz darauf eingegangen, inwiefern bereits ähnliche Studien existieren. Da es nach Wissen der Autorin kaum bzw. keine Studien gibt, die diese drei Forschungsstränge miteinander verknüpfen, erfolgt ein ausführlicher Forschungsstand zu den einzelnen Strängen in den jeweiligen theoretischen Kapiteln.

Generell lässt sich festhalten, dass die Arbeit eine sozialkonstruktivistische Perspektive einnimmt. Demnach grenzt sie sich in Bezug auf die Forschung zu Messinstrumenten von denjenigen Untersuchungen ab, in denen im positivistischen Sinne Instrumente in Bezug auf ihre Validität und Reliabilität analysiert werden (z. B. Gärtner und Pant 2011). Dagegen knüpft diese Arbeit in Bezug auf Messinstrumente an **bewertungssoziologische Überlegungen** (vgl. Lamont 2012)

sowie Studien zur Entwicklung von Indikatoren (z. B. Desrosières 2000) an. Diesen ist gemein, dass die Instrumente bzw. die Präsentation der Daten, die mithilfe der Instrumente gesammelt werden, nicht als objektiv angesehen werden und der situative Kontext von Bewertungsprozessen zu berücksichtigen ist. In Bezug auf Qualitätsurteile sind insbesondere die Arbeiten von Boltanski und Thévenot (2007) zu nennen. Diese haben den „Économie-de-conventions“-Ansatz (EC) entwickelt und argumentieren, dass Akteure in unsicheren Situationen, z. B. bei dem Entstehen von Qualitätsurteilen, auf bestimmte Konventionen zurückgreifen, um ihr Handeln zu legitimieren. Daran anknüpfend geht es im Rahmen dieser Arbeit um die Frage, wie die Entwickler_innen der Instrumente ihr Vorgehen begründen und auf welche Werteordnungen hierbei zugegriffen wird. In Bezug auf Praktiken der Bewertung geht es daher im Fokus nicht darum „aufzuzeigen, wie diese Interaktion [bei der Entwicklung der Instrumente, ap] beschaffen sein sollte, sondern wie sie – historisch und sozial – gewesen ist“ (Desrosières 2005, S. 4).

Des Weiteren knüpft die Arbeit an erziehungswissenschaftliche und soziologische Untersuchungen zur **Implementation von politischen Vorgaben**, z. B. Gesetzen, an. Hierbei wird in den Erziehungswissenschaften im deutschsprachigen Raum häufig der Rekontextualisierungsansatz von Fend (2008) angewandt (z. B. Amrhein 2011). Im Rahmen dieser Arbeit wird auf den Sensemaking-Ansatz in der adaptierten Fassung von Spillane et al. (2002) zurückgegriffen. Dieser weist sehr viele Parallelen zum Fend'schen Rekontextualisierungsansatz auf, aber stellt insofern eine Erweiterung dar, als dass der Ansatz die Komplexität von kognitiven Deutungsprozessen berücksichtigt und dabei unter anderem emotionale Aspekte auf individueller Ebene sowie gesellschaftliche Strukturen in den Blick nimmt. Des Weiteren gibt es Anknüpfungspunkte zwischen dem Ansatz des Sensemakings und dem Educational-Governance-Ansatz (siehe z. B. Altrichter et al. 2007; Dietrich 2018): So ist beispielsweise beiden Ansätzen gemein, dass nicht davon ausgegangen wird, dass formal bestehende Hierarchieverhältnisse handlungsleitend seien (nähere Ausführung → Kapitel 3).

Während beim Sensemaking-Ansatz häufig die Implementation von konkreten Maßnahmen näher untersucht wird, wird im Rahmen der Arbeit nicht primär betrachtet, inwiefern Art. 24 BRK in den Instrumenten der Schulinspektionen implementiert wurde. Stattdessen wird Inklusion als eine politische Vorgabe verstanden, die in verschiedenen rechtlichen Konventionen der UN verankert ist, z. B. auch in Art. 13 des UN-Sozialpakts. Damit wird das Verständnis von Inklusion, welches in der BRK beschrieben wird, miteingeschlossen, aber weitere Aspekte, z. B. andere Personengruppen, können dadurch in den Blick genommen werden. Daher wird im Folgenden Inklusion allgemein als politische Vorgabe beschrieben.

Des Weiteren lassen sich Bezüge **zur Inklusionsforschung** herstellen: So wird im ersten Teil der empirischen Analyse untersucht, inwiefern ein bestimmtes Verständnis von Inklusion in den Instrumenten berücksichtigt wird. Dabei grenzt sich diese Arbeit von Studien ab, die unter Inklusion nur den gemeinsamen Unterricht von Schüler_innen mit und ohne Behinderung verstehen (z. B. Bertelsmann-Stiftung 2014). So wird im Folgenden ein weiter Inklusionsbegriff (vgl. Ainscow et al. 2006) verwendet, d. h. alle Schüler_innen werden in ihrer Individualität gesehen und verschiedene Heterogenitätsdimensionen, die in Bezug auf Teilhabe und Diskriminierung von Relevanz sind, können in den Blick genommen werden (Booth und Ainscow 2002; Prengel 2006; siehe auch Piezunka 2015). Hierbei wird unter anderem auf die theoretischen Überlegungen von Hinz (2002), Ainscow et al. (2006) und Prengel (2006) zurückgegriffen. Im zweiten Teil der Analyse wird untersucht, wie Inklusion – wenn der Begriff explizit verwendet wird – von Seiten der Schulinspektionen gedeutet wird. Hier knüpft die Arbeit an bestehende Untersuchungen an, die sich damit auseinandersetzen, inwiefern es unterschiedliche Deutungen

von Inklusion gibt (z. B. Piezunka et al. 2017; siehe auch Kruse und Dederig 2018) bzw. wie Inklusion in den einzelnen Ebenen des Bildungssystems implementiert wurde (z. B. Moser und Egger 2017).

1.4 Aufbau der Arbeit

Nach der Einleitung beginnt das zweite Kapitel mit einer Darstellung dessen, welche Verständnisse von Inklusion den deutschsprachigen Diskurs prägen und worin sie sich unterscheiden. Um im ersten Teil der Analyse eine normative Vergleichsfolie nutzen zu können, ist es – wie oben bereits dargestellt – notwendig, den Fokus auf ein Verständnis von Inklusion zu legen, welches in der empirischen Analyse als Kategoriensystem genutzt werden kann. Zusätzlich enthält das zweite Kapitel Informationen über den aktuellen Stand der Umsetzung von Inklusion in Deutschland bzw. zur vorhandenen Datenlage.

Im dritten Kapitel folgt eine Beschreibung dessen, welche Funktionen Schulinspektionen zugeschrieben werden. Des Weiteren wird dargestellt, wo sich diese im deutschen Bildungssystem verorten lassen. Hierfür dient der Educational-Governance-Ansatz als sensibilisierende Perspektive, um Akteurskonstellationen und Abhängigkeiten abbilden zu können.

Im vierten Kapitel wird beschrieben, inwiefern Messinstrumente dazu dienen, Objekte zu klassifizieren und welche Wirkung mit bestimmten Indikatoren einhergeht. Dabei wird auf theoretische Überlegungen aus der Bewertungssoziologie zurückgegriffen. Zusätzlich wird im Rahmen eines Exkurses herausgearbeitet, dass es keinen Konsens darüber gibt, was unter einer guten Schule zu verstehen ist und welche Verständnisse im deutschsprachigen Raum dominieren.

Im fünften Kapitel folgt die Darstellung des Sensemaking-Ansatzes: Um die Deutung von Inklusion und die Entwicklung von Instrumenten zu verstehen, wird auf den Sensemaking-Ansatz von Weick (1985) bzw. auf die adaptierte Fassung von Spillane et al. (2002) zurückgegriffen. Dieser Ansatz kommt aus der Organisationssoziologie und trifft Annahmen darüber, welche Dimensionen bei kognitiven Deutungsprozessen eine Rolle spielen. Dabei geht es um Deutungsprozesse von Objekten, Konzepten oder Situationen, die für die Sensemaker unbekannt bzw. mehrdeutig sind.

Im sechsten Kapitel folgt eine Übersicht über das empirische Design der Arbeit. Die Arbeit besteht aus zwei empirischen Teilen: der Analyse der Bewertungsinstrumente (12 Fälle) sowie Experteninterviews (18 Interviews in 9 Fällen). Die empirische Analyse der Bewertungsinstrumente konzentriert sich auf den Zeitraum nach 2009, da in diesem Jahr die BRK von Deutschland ratifiziert wurde und ab diesem Zeitpunkt Inklusion als politische Vorgabe im öffentlichen Raum intensiver diskutiert wurde. Im Rahmen der Arbeit werden die Bewertungsinstrumente analysiert, die im Oktober 2015 von den Schulinspektionen genutzt wurden. Dieser Zeitpunkt bietet sich insofern an, da die Bewertungsinstrumente in fast allen Bundesländern überarbeitet wurden, nachdem alle Schulen einmal evaluiert worden waren. Die Überarbeitung der Instrumente fiel dabei in den Zeitraum nach der Ratifizierung der BRK. Um die Entwicklung der Instrumente abbilden zu können, wurden Interviews mit denjenigen Mitarbeiter_innen der Schulinspektionen geführt, die maßgeblich an der Entstehung der Instrumente beteiligt waren (im Folgenden „Instrumenteentwickler_innen“).

Durch die Analyse der Messinstrumente kann abgebildet werden, inwiefern Inklusion in den Messinstrumenten berücksichtigt wird. Durch die Analyse der Interviews mit Instrumenteentwickler_innen lässt sich herausarbeiten, wie sich diese im Rahmen von Aushandlungsprozessen

mit ihren Kolleg_innen sowie auf individueller Ebene mit Inklusion auseinandergesetzt haben und welche Rahmenbedingungen hierbei prägend wirkten. Dabei wird das gesamte empirische Material mit Hilfe der qualitativen Inhaltsanalyse nach Gläser und Laudel (2009) ausgewertet. Ausgehend von den unterschiedlichen Fragestellungen wird bei der Analyse auf verschiedene Kategoriensysteme zurückgegriffen.

Nach der Fallbeschreibung folgt im siebten Kapitel die Darstellung der empirischen Ergebnisse bezüglich der Frage, ob und wie Inklusion von Seiten der Schulinspektionen berücksichtigt wurde. Im ersten Schritt wird näher betrachtet, inwiefern bestimmte Konzepte, die auf dem oben bereits erwähnten menschenrechtlichen Inklusionsverständnis basieren, in den Instrumenten der bundesländerspezifischen Schulinspektionen sichtbar werden. Dadurch lässt sich abbilden, ob Inklusion als ein Teilaspekt von Schulqualität in den Instrumenten thematisiert wird und welche Relevanz diesem beigemessen wird. Dabei werden die verfügbaren Instrumente als Ganzes analysiert – unabhängig davon, ob der Inklusionsbegriff explizit aufgeführt wird oder nicht. Im zweiten Schritt wird analysiert, wie Inklusion als Konzept von Seiten der Schulinspektionen gedeutet wird. Um zu analysieren, was die Schulinspektionen in den Messinstrumenten unter Inklusion verstehen, sind nur diejenigen Instrumente relevant, die von Seiten der Schulinspektion als „inklusionsorientiert“ gekennzeichnet wurden. Im dritten Schritt folgt auf der Grundlage der Experteninterviews eine Analyse, wie die Instrumenteentwickler_innen als Einzelpersonen Inklusion gedeutet haben.

Im achten Kapitel wird dargestellt, welche Rahmenbedingungen die Entwicklung der Instrumente in Bezug auf Inklusion geprägt haben. Dabei zeigt sich, dass eine Vielzahl von unterschiedlichen Bedingungen eine Rolle gespielt hat. Im neunten Kapitel, der Diskussion, werden die Ergebnisse der Arbeit nochmal zusammengefasst und Bezüge zu bereits vorhandenen Studien hergestellt sowie Implikationen für die Praxis beschrieben. Im zehnten Kapitel folgt ein abschließendes Fazit und ein Ausblick auf zukünftige Forschung.

2 Schulische Inklusion – eine Begriffsklärung

Im folgenden Kapitel wird zunächst ein Überblick darüber gegeben, welche Verständnisse von schulischer Inklusion existieren und worin sie sich unterscheiden. Im Rahmen des zweiten Abschnitts (→ 2.2) wird ein Verständnis von Inklusion näher beschrieben, welches in der empirischen Analyse als normative Vergleichsfolie dienen wird. Damit ist es möglich, die Inklusionsverständnisse der Schulinspektionen einzuordnen und zu diskutieren. Die Festlegung auf ein Inklusionsverständnis ermöglicht es darüber hinaus, zu untersuchen, ob in den Instrumenten Aspekte dieses Verständnisses operationalisiert wurden, ohne dass Inklusion als Begriff explizit verwendet wird.

Abschließend wird im dritten Abschnitt (→ 2.3) dargestellt, wie der aktuelle Stand der Umsetzung von Inklusion in Deutschland ist. Dies dient als Kontextinformation, um den situativen Kontext von Schulinspektionen einordnen zu können. Hierbei wird auf Studien aus der empirischen Bildungsforschung zurückgegriffen. Jedoch lassen sich aktuell nur begrenzt Aussagen über die Umsetzung von Inklusion treffen, was unter anderem darauf zurückgeführt werden kann, dass es an einem umfangreichen Set an Indikatoren fehlt, anhand dessen die Umsetzung von Inklusion gemessen werden kann.

2.1 Unterschiedliche Verständnisse von Inklusion

In den Erziehungswissenschaften hat Inklusion als Begriff in den letzten Jahren durch die Ratifizierung der BRK an Relevanz gewonnen. Es gibt jedoch keine „allgemein anerkannte Definition von Inklusion“ (Moser 2012b, S. 24; Grosche 2015, siehe auch Campbell 2002, S. 12; Heinrich et al. 2013, S. 69). Demnach existieren in der Literatur unterschiedliche Verständnisse von Inklusion (z. B. Wocken 2009; Hinz 2004; Degener und Mogge-Grotjahn 2012; Ainscow et al. 2006, S. 14; Deutsche UNESCO-Kommission 2014; Biewer 2009). Dabei lassen sich diese unter anderem auf die jeweiligen theoretischen Rahmungen zurückführen. So dient Inklusion beispielsweise in der Systemtheorie nach Luhmann (1995) als deskriptive Kategorie, um die Teilhabe an Funktionssystemen zu beschreiben. Für Luhmann ging mit dem Konzept (zunächst) keine normative Positionierung einher.¹ Dagegen dient Inklusion als normative Kategorie, wenn es unter anderem bei Degener und Mogge-Grotjahn (2012, S. 12) als Menschenrecht beschrieben wird. Die jeweiligen Systembezüge werden auch im Hinblick auf konkrete Aspekte sichtbar: Beispielsweise erfordert eine rechtswissenschaftliche Perspektive auf Inklusion, dass diese umsetzbar sei. Wohingegen eine pädagogische Perspektive impliziert, dass Inklusion nicht getrennt von bestehenden gesellschaftlichen Funktionen von Schule betrachtet werden kann. Anhand dieser Beispiele wird deutlich, dass die Verortung des Inklusionsbegriffs in bestehende theoretische Rahmungen mit bestimmten Implikationen einhergeht. Durch die Anwendung des Inklusionsbegriffs im Rahmen der jeweiligen Betrachtungsweisen werden daher stets verschiedene Aspekte sichtbar: „So ist es möglich, dass Inklusion als Leitidee, als Programm,

¹ Hier folgte bei Luhmann ein Wandel der Argumentation am Ende seines Lebens. So erhielt er einen Einblick in das Leben in brasilianischen Favelas. Dort wurde ihm deutlich, welche Implikationen damit einhergehen, wenn ein Mensch aus allen Funktionssystemen exkludiert wird.

als Ziel, als Vision, als Menschenrecht, zur Systembeschreibung, als Wert, als Handlungspraxis beschrieben wird“ (Neißl 2016, S. 1).

Unabhängig von den jeweiligen Systembezügen stellt sich die Frage, welche Verständnisse von schulischer Inklusion im Diskurs existieren und worin sie sich unterscheiden. Verschiedene Autor_innen (z. B. Göransson und Nilholm 2014; Ainscow et al. 2006; Kruse und Dederling 2018; Piezunka et al. 2017) haben dies systematisch analysiert: So hat beispielsweise Grosche (2015, 33f.) durch die Auswertung von deutschsprachigen Publikationen 28 Beschreibungen von schulischer Inklusion identifiziert. Durch die Analyse von englischsprachigen Datenbanken aus dem Bereich der Bildungsforschung konnten Göransson und Nilholm (2014) vier Verständnisse von Inklusion herausarbeiten. Im Vergleich zu den Beschreibungen bei Grosche sind diese abstrakter formuliert. Seine Beschreibungen lassen sich unter den von Göransson und Nilholm genannten Verständnissen subsumieren.

In Bezug auf diese Arbeit sind die oben aufgeführten Analysen insofern von Relevanz, als dass sie zum einen bestätigen, dass es kein allgemeingültiges Verständnis von Inklusion gibt. Des Weiteren geben sie Hinweise darauf, welche Dimensionen von Relevanz sind, um die einzelnen Inklusionsverständnisse voneinander zu unterscheiden (vgl. Piezunka et al. 2017):

- Die jeweiligen Inklusionsverständnisse sind auf **unterschiedlichen Ebenen des Bildungssystems**, z. B. Unterrichtsebene, anzusiedeln.
- Als **Ziel von Inklusion** kann der Zugang zur Regelschule gesehen werden, jedoch gibt es auch Verständnisse, wonach der Zugang zur Regelschule nur eine von vielen Maßnahmen ist, um gesellschaftliche Teilhabe zu ermöglichen. Bei letzterem ist daher das Ziel von Inklusion die gesellschaftliche Teilhabe aller Schüler_innen.
- Auch bei der **Begründung** dafür, warum Inklusion notwendig sei, wird nach Felder (2012) auf unterschiedliche Rechtfertigungsprinzipien zurückgegriffen, z. B. auf die gesetzliche Verankerung von Inklusion. Des Weiteren wird die Umsetzung von Inklusion – verstanden als gemeinsamer Unterricht von Schüler_innen mit und ohne Behinderung – auch damit begründet, dass dadurch alle Kinder und Jugendliche in Bezug auf den Lernprozess profitieren würden. Diese Form der Legitimation kann jedoch nach Hinz und Boban (2017) problematisch sein: Dies ist beispielsweise der Fall, wenn die Integration von Schüler_innen mit Behinderung nicht dazu führt, dass alle Schüler_innen davon profitieren. Basierend auf dieser Begründung, würde dann die Umsetzung von Inklusion ihre Legitimation verlieren. Hinz und Boban (2017) grenzen sich daher von jener Legitimation ab und argumentieren aus einer menschenrechtlichen Perspektive, dass die Situation von Kindern, deren Teilhabe in Gefahr ist, im Mittelpunkt stehen sollte. Zusammenfassend zeigt sich, dass auf unterschiedliche Argumente Bezug genommen wird, um die Umsetzung von Inklusion zu begründen. Felder kritisiert jedoch, dass es häufig auch an Rechtfertigungen fehle: „Vielmehr dient Inklusion als – oft unhinterfragter – Leitbegriff und Horizont für einen anzustrebenden Zustand lebensweltlicher Eingebundenheit“ (Felder 2012, S. 18).
- Des Weiteren unterscheiden sich die Inklusionsverständnisse in der Frage, wer **die Zielgruppe von Inklusion** sei. Häufig wird zwischen einem engen Inklusionsbegriff unterschieden, der Menschen mit zugeschriebener Behinderung in den Blick nimmt sowie einem weiten Inklusionsbegriff, der alle Schüler_innen umfasst, aber meistens einen besonderen Fokus auf bestimmte Differenzlinien legt (vgl. Ainscow et al. 2006, S. 16). In dem Zusammenhang werden häufig Differenzlinien aufgeführt, die in Bezug auf Teilhabe am Bildungssystem von Relevanz sind (nähere Ausführung folgt → Abschnitt 2.2.4).

- Auch die Frage der **Realisierbarkeit** spielt eine Rolle, um die einzelnen Inklusionsverständnisse voneinander zu unterscheiden: Aus juristischer Perspektive müssen bestimmte Anforderungen direkt umgesetzt werden und Inklusion als Ganzes kann progressiv implementiert werden (vgl. Piezunka et al. 2018). Dagegen gibt es andere Verständnisse, wonach bei der Beschreibung von Inklusion auf die Metapher des Nordsterns zurückgegriffen wird (vgl. Hinz 2010). Danach wird Inklusion als Prozess verstanden und die Frage der Realisierbarkeit verliert insofern an Relevanz, als dass es der ständigen Reflektion darüber bedarf, inwiefern Strukturen und Praktiken weiterhin inklusiv sind.

Die verschiedenen Ausprägungen innerhalb der einzelnen Dimensionen machen nochmals deutlich, dass es unterschiedliche Verständnisse von Inklusion gibt. Daran anknüpfend haben sich Piezunka et al. (2017) mit der Frage auseinandergesetzt, inwiefern sich ein gemeinsamer Kern herausarbeiten lässt. Hierfür haben sie Inklusionsforscher_innen befragt und insgesamt vier verschiedene Verständnisse von Inklusion herausgearbeitet. Im Rahmen der Analyse konnten sie folgenden Kern identifizieren, der auf alle von ihnen herausgearbeiteten Inklusionsverständnisse zutrifft: **die Überwindung von Diskriminierung aufgrund von sozial konstruierten Differenzlinien** (vgl. ebd.). Dieser Kern weist jedoch in den einzelnen Verständnissen unterschiedliche Ausprägungen auf: Beispielsweise variiert, was unter Überwindung von Diskriminierung verstanden wird bzw. welche Differenzlinien von Relevanz sind (vgl. ebd.).

Abschließend lässt sich festhalten, dass es an einer allgemein anerkannten Definition von Inklusion im schulischen Kontext fehlt und diverse Verständnissen im deutschsprachigen Diskurs gegenwärtig existieren. In Bezug auf die Fragestellung der Arbeit bedeutet das Fehlen einer allgemeingültigen Definition von Inklusion, dass nicht davon ausgegangen werden kann, dass alle Schulinspektionen das Gleiche meinen, wenn sie in den Bewertungsinstrumenten Inklusion als Begriff verwenden. Daher stellt die Untersuchung davon, was die Schulinspektionen in ihren Instrumenten unter Inklusion verstehen, einen Teilaspekt der empirischen Analyse dar (nähere Ausführung → Kapitel 6). Hierfür dienen die oben beschriebenen Dimensionen als Orientierung. Bei der ersten Strukturierung des Materials und in der Analyse können weitere Dimensionen ergänzt werden. Wie bereits oben erwähnt ist eine normative Vergleichsfolie notwendig, um die jeweiligen Deutungen einordnen zu können. So argumentiert Felder, dass durch die Existenz von unterschiedlichen Inklusionsverständnissen die Gefahr besteht, dass Inklusion „als Containerbegriff für eine Vielzahl von pädagogischen Konzepten und Zugängen fungiert“ (Felder 2012, S. 18). Das Fehlen eines allgemeingültigen Verständnisses von Inklusion wird im Hinblick auf die empirische Analyse problematisch, wenn es darum geht zu beurteilen, ob bestimmte Instrumente „inklusionsorientiert“ sind oder nicht. Demnach wären im nächsten Schritt auch keine Aussagen darüber möglich, ob Inklusion als politische Vorgabe umgesetzt ist oder nicht bzw. welche Implikationen sich aus der Umsetzung ergeben. Zusätzlich erlaubt die Festlegung auf eine normative Vergleichsfolie zu analysieren, inwiefern Teilaspekte von Inklusion in den Instrumenten berücksichtigt werden ohne dass sie explizit mit dem Inklusionsbegriff verknüpft werden. Daher werden im Rahmen der Arbeit bestimmte Prämissen festgelegt, die als normative Vergleichsfolie dienen. Diese werden im Folgenden näher beschrieben und hierbei wird zugleich begründet, weshalb auf dieses Verständnis von Inklusion bzw. auf die dazugehörigen Konzepte zurückgegriffen wird.

2.2 Menschenrechtliches Verständnis

Im Rahmen der Arbeit dient ein Verständnis von Inklusion als Orientierung, welches auf den folgenden Prämissen basiert:

- **Durch Inklusion werden im schulischen Setting die folgenden Ziele angestrebt: individuell bestmögliche Qualifizierung; politische Partizipation sowie das Erfahren von intersubjektiver Anerkennung.**
- **Inklusion zielt auf alle Schüler_innen ab. Bei der empirischen Analyse wird im Hinblick auf die Verwendung von sozial konstruierten Differenzlinien eine kritische Perspektive eingenommen.**
- **Das Inklusionsverständnis beschränkt sich auf die Einzelschulebene bzw. darunter liegende Ebenen.**

Wie lässt es sich rechtfertigen, dass bei der Vielzahl an unterschiedlichen Inklusionsverständnissen diese Prämissen gesetzt wurden? Inhaltliche Hinweise ergeben sich zunächst aus der Fragestellung dieser Arbeit: Da es um die Implementation von Inklusion als politische Vorgabe in den Bewertungsinstrumenten der Schulinspektionen geht, ergibt sich die Notwendigkeit für einen normativ besetzten Inklusionsbegriff in Abgrenzung zu Verständnissen von Inklusion als deskriptive Kategorie. In Bezug auf den Gegenstand der Arbeit ergeben sich zwei weitere Setzungen: So lässt sich der Fokus auf die Einzelschulebene durch das Handlungsfeld von Schulinspektionen erklären (nähere Ausführung → Abschnitt 2.2.2). Des Weiteren wird im Rahmen dieser Arbeit an den Inklusionsbegriff angeknüpft, der in der BRK und anderen Dokumenten der UN verwendet wird. So geht es im Rahmen der Arbeit um die Implementation von Inklusion als politische Vorgabe auf der Ebene der Bildungsadministration (nähere Ausführung → Abschnitt 2.2.1).

Bei der Konkretisierung der oben beschriebenen Prämissen wird auf die bestehende wissenschaftliche Literatur zurückgegriffen, um die von der Autorin getroffenen Setzungen zu erklären und zu begründen. Zum Teil basieren diese Setzungen jedoch auf Wertüberzeugungen, z. B. was als primäre Funktionen von Schule angesehen wird. Da Wertüberzeugungen subjektiv sind, ist es möglich, auch andere als die von der Autorin vorgenommenen Setzungen zu treffen und somit auf abweichende Verständnisse bzw. Konkretisierungen zurückzugreifen. Eine objektive Letztbegründung ist daher nicht möglich. Darüber hinaus stellen die hier verwendeten Konzepte in Bezug auf die bestehende wissenschaftliche Literatur keine Neuentwicklung dar, sondern können als Konkretisierung von bestehenden Verständnissen gelesen werden (z. B. Biewer 2009, S. 193) bzw. weisen hohe Überschneidungen auf, z. B. mit der Definition von Ainscow et al.: „The processes of increasing the participation of students in, and reducing their exclusion from, the curricula, cultures and communities of local schools“ (Ainscow et al. 2006, S. 25). Daher werden die Bezüge zu diesen theoretischen Arbeiten im Rahmen des Kapitels immer wieder hergestellt.

2.2.1 Menschenrechtliche Perspektive

Die Analyse der Bewertungsinstrumente konzentriert sich auf den Zeitraum nach der Ratifizierung der BRK. Dies basiert auf der Annahme, dass die Ratifizierung der BRK im Fall von Deutschland die Auseinandersetzung mit dem Inklusionsbegriff in der breiten Öffentlichkeit (vgl. Werning 2014) und somit auch bei den Schulinspektionen maßgeblich initiiert hat.

Im Folgenden wird argumentiert, dass durch Inklusion der Schutz von bestimmten menschenrechtlichen Konzepten angestrebt wird (vgl. Kruschel 2017). So werden im Zusammenhang mit Inklusion – sowohl in den Menschenrechtskonventionen der UN als auch in der wissenschaftlichen Literatur – die folgenden zwei Konzepte immer wieder genannt: die **Überwindung von Diskriminierung** und das **Ermöglichen von Teilhabe** (z. B. Degener und Mogge-Grotjahn 2012, S. 65; c. Lindmeier und B. Lindmeier 2015, S. 43; Boban und Hinz 2003; Werning 2014; Ainscow et al. 2006, S. 18; Prengel 2006; Boban und Hinz 2003; Alicke und Linz-Dinchel 2012; Moser 2017; Graumann 2012; Reitz 2015; Rudolf 2017). Im Folgenden werden diese beiden menschenrechtlichen Konzepte als Ziele von Inklusion verstanden. Es bedarf jedoch einer Konkretisierung sowie Festlegung auf bestimmte Lesarten. So gibt es beispielsweise im Fall von Teilhabe kein allgemeingültiges Verständnis, was darunter zu verstehen ist bzw. wird es häufig synonym mit anderen Konzepten wie Partizipation verwendet.

Bei der Festlegung auf ein Inklusionsverständnis ist darüber hinaus zwischen der Zielsetzung der BRK sowie dem in der BRK verwendeten Inklusionsbegriff zu unterscheiden. So ist die BRK nur eine von vielen Konventionen und Empfehlungen der UN, die den Inklusionsbegriff verwendet. In der BRK geht es um die Rechte von Menschen mit diagnostizierten körperlichen und psychischen Beeinträchtigungen, die aufgrund der bestehenden Strukturen in ihrer Teilhabe behindert werden. Aus der BRK ergeben sich Anforderungen, die die Teilhabe dieser Gruppe gewährleisten. Dies impliziert jedoch nicht, dass sich Inklusion als Begriff nur auf Menschen mit zugeschriebener Behinderung bezieht (vgl. Reitz 2015, S. 6; Rudolf 2017, S. 36), auch wenn dies in der öffentlichen Debatte häufig so diskutiert wird. Inklusion bezieht sich auf die Teilhabe aller Menschen. Dies wird deutlich, wenn weitere Konventionen und Empfehlungen der UN als Kontext der BRK betrachtet werden. So zeichnet Biewer (2009, 128f.) nach, wie sich die UN im Laufe der Zeit von einem engen Inklusionsbegriff verabschiedet hat und nun verstärkt verschiedene Differenzlinien in den Blick nimmt. So wird Inklusion in dem Referenzdokument, das anlässlich der „International Conference on Education“ der UNESCO (2008 in Geneva) verteilt wurde, folgendermaßen definiert:

„Inclusive education is a process that involves the transformation of schools and other centres of learning so as to cater for all children – including boys and girls, students from ethnic minorities, those affected by HIV and AIDS, and those with disabilities and learning difficulties“ (vgl. UNESCO 2008, S. 5; siehe auch UNESCO 2017b).

Des Weiteren wird in den Sustainable Development Goals (SDG), die im September 2015 verabschiedet wurden, folgendes Ziel formuliert: „Ensure inclusive and quality education for all“ (UNESCO 2017a). Dabei geht es um die Teilhabe aller Schüler_innen und eine Vielzahl an Differenzlinien, die in Bezug auf Teilhabe und Diskriminierung von Relevanz sind, werden explizit benannt, z. B. Gender, Kinder in belastenden Situationen, Disability, usw. Demnach soll der Zugang zu Bildung für alle gewährleistet werden. Daran anknüpfend werden als Zielgruppe von Inklusion im Folgenden alle Schüler_innen verstanden (nähere Ausführung → Abschnitt 2.2.4).

2.2.2 Fokus auf die Einzelschulebene

Der Fokus auf die Einzelschulebene ergibt sich aus der Fragestellung der Arbeit. So geht es im Rahmen der Arbeit um die Messung von Qualität an einzelnen Schulen. Demnach entwickeln Schulinspektionen Messinstrumente, die sich auf diese Handlungsebene bzw. darunterliegende Ebenen, wie Unterricht, konzentrieren. Durch den Gegenstand der Arbeit ergibt sich daher die

Anforderung, dass nur von Relevanz ist, was von der Einzelschule gestaltbar bzw. veränderbar ist. Hierbei ist jedoch zu berücksichtigen, dass nur begrenzt festlegbar ist, wo der Handlungsbereich der Einzelschule anfängt bzw. wieder aufhört. So könnte beispielsweise argumentiert werden, dass der Unterricht nicht beurteilbar sei, weil die Schule nur begrenzt Einfluss auf die Auswahl von Lehrkräften habe.

Der Fokus auf die Einzelschulebene impliziert jedoch nicht, dass dichotome Aussagen darüber möglich sind, ob eine Schule als Gesamteinheit inklusiv sei. Schulen sind komplexe Organisationsseinheiten. So kann es beispielsweise sein, dass bestimmte Aspekte der Unterrichtsgestaltung inklusiv sind, aber die organisatorischen Strukturen nicht. Darüber hinaus werden im schulischen Alltag immer wieder Entscheidungen getroffen, die die Gefahr bergen einzelne Menschen zu exkludieren. Daher ist ein einmal formulierter Status über eine Einzelschule wenig aussagekräftig. Stattdessen bietet es sich eher an, Aussagen darüber treffen, inwiefern an einer Schule Gelingensbedingungen im Hinblick auf die Umsetzung von Inklusion vorhanden sind.

2.2.3 Teilhabe und Nichtdiskriminierung in Bezug auf Schule

Die oben genannten Menschenrechte – Teilhabe und Überwindung von Diskriminierung – geben einen ersten Rahmen vor. Jedoch fehlt es in bestehenden UN-Konventionen wie der BRK an einer konkreteren Beschreibung davon, was diese im Hinblick auf den schulischen Kontext implizieren bzw. was es für den deutschen Fall bedeutet. Hierfür wird im Folgenden auf bestehende Arbeiten aus der Schultheorie zurückgegriffen. Wie bereits oben erwähnt, ist es hierbei notwendig normative Setzungen vorzunehmen, z. B. was relevante Funktionen von Schulen sind. Dies hat zur Folge, dass das hier dargestellte Verständnis nur eine mögliche Interpretation für das schulische Setting in Deutschland darstellt. Im Folgenden werden die Menschenrechte „Teilhabe“ und „Überwindung von Diskriminierung“ in Bezug auf die Einzelschulebene konkretisiert.

Konkretisierung des Teilhabebegriffs in Bezug auf den schulischen Kontext: Für die Konkretisierung des Teilhabebegriffs wird auf die gesellschaftlichen Funktionen von Schule nach Fend (vgl. Fend 2006; siehe auch Biesta 2009) zurückgegriffen:

Qualifikationsfunktion: Die Schule wird heute (...) als wesentlich für die Erzeugung eines <Humankapitals> im Sinne der Berufsfähigkeit der jüngeren Generation gesehen (...)“ (Fend 2006, S. 52). Schule hat somit die Aufgabe bestimmte Fähigkeiten und Wissen zu vermitteln, damit Schüler_innen am gesellschaftlichen Leben teilhaben können.

Enkulturationsfunktion: „Der Erwerb von Fähigkeiten der Reflexion, der eigenständigen Urteilsbildung und der moralischen Entscheidungsfindung“ (ebd., S. 49). So hat Schule die Aufgabe bestimmte Werte und Normen zu vermitteln.

Integrationsfunktion: „Soziale Identitätsbildung, Identifikation und soziale Bindung als Grundlage für soziale Verantwortung“ (ebd., S. 53).

Allokationsfunktion: Verschiedene Formen der Selektion, z. B. Bewertung durch Noten oder Verteilung auf Schulformen, die im schulischen Setting stattfindet, leisten einen Beitrag dazu, dass einzelne Personen auf soziale Positionen innerhalb der Gesellschaft verteilt werden.

Unter Berücksichtigung von drei Funktionen: der Qualifikations-, Enkulturations- und Integrationsfunktion lassen sich zwei Aspekte herausarbeiten (Erläuterung zur Allokationsfunktion

→ Abschnitt 2.2.4): Ausgehend von der Qualifikationsfunktion wird Schule als ein Ort gesehen, der zur gesellschaftlichen Teilhabe befähigt, z. B. indem Schüler_innen dazu qualifiziert werden, im Anschluss an den Schulbesuch eine berufliche Ausbildung oder ein Studium zu absolvieren. Ausgehend von der Integrations²- und Enkulturationsfunktion kann Teilhabe auch als (politische) Partizipation in einer demokratischen Gesellschaft verstanden werden. Hiermit ist gemeint, dass Schüler_innen die Möglichkeit haben den schulischen Alltag und Strukturen mitzubestimmen und gestalten.

„Sie umfasst neben politischer Teilhabe auch die Übernahme von Verantwortung für das Gemeinsame durch aktive Mitgestaltung und impliziert sozialen Zusammenschluss, Kooperation und Aushandlungsprozesse mit anderen Menschen und heterogenen Gruppen“ (Eikel 2006, S. 6).

Dahinter steht der Gedanke, dass Schüler_innen durch politische Partizipation demokratische Kompetenzen erwerben, z. B. Urteils- und Entscheidungsfähigkeit (ebd., S. 10). Ausgehend von der Integrations- und Enkulturationsfunktion wird im Rahmen dieser Arbeit in Bezug auf Partizipation primär auf eine demokratiepädagogische und nicht auf eine lerntheoretische³ Argumentation zurückgegriffen. Demnach wird das Ermöglichen von politischer Partizipation im schulischen Kontext als notwendige Voraussetzung für das Gelingen einer Demokratie gesehen (z. B. Moldenhauer 2015, S. 21). In Bezug auf die bestehende demokratiepädagogische Literatur gibt es kein einheitliches Verständnis davon, wodurch sich politische Partizipation in pädagogischen Settings kennzeichnet (vgl. Deutsche Gesellschaft für Demokratiepädagogik e.V. 2014, S. 14). Des Weiteren ist zu berücksichtigen, dass bei der Frage nach möglicher Mitgestaltung von Schüler_innen im schulischen Alltag Grenzen existieren: Beispielsweise gibt es bestimmte gesetzliche Vorgaben wie die Schulpflicht, die nicht von Seiten der Schüler_innen verändert werden können (vgl. Moldenhauer 2015, S. 9). Innerhalb dieser Grenzen ist das Ermöglichen von Partizipation jedoch ein notwendiges Instrument für die Herausbildung von autonomen und selbstbestimmten Subjekten (vgl. ebd.). Aus menschenrechtlicher Perspektive ergibt sich demnach ein „subjektorientiertes und rechtebasiertes Verständnis von Partizipation“ (vgl. Reitz 2015, S. 5), das auch in der Kinderrechtskonvention (KRK) (vgl. Netzwerk Menschenrechte 2017; siehe auch Deutsche Gesellschaft für Demokratiepädagogik e.V. 2014) verankert ist. Zusammenfassend lassen sich demnach zwei Verständnisse von Teilhabe ableiten: Teilhabe als Qualifizierung von Schüler_innen sowie im Sinne politischer Partizipation.

Konkretisierung des Konzepts „Überwindung von Diskriminierung“ in Bezug auf den schulischen Kontext: Als weiteres Ziel von Inklusion wird im Rahmen dieser Arbeit die Überwindung von Diskriminierung gesehen. In diesem Zusammenhang ließe sich argumentieren, dass das Ermöglichen von Teilhabe ein Mittel ist, um Diskriminierung zu überwinden. Diskriminierung kann jedoch weiter gefasst werden. So argumentiert Rudolf: „Teilhabe ist nur möglich, wenn keine Diskriminierung stattfindet“ (Rudolf 2017, S. 15). Demnach können darunter auch verletzendende Äußerungen in persönlichen Interaktionen fallen, z. B. stereotype Zuschreibungen, die sich negativ auf die Identitätsbildung des Individuums auswirken können

2 Weitere Aspekte, die in Bezug auf die Integrationsfunktion von Relevanz sind, wurden bei der Konkretisierung des Konzepts „Überwindung von Diskriminierung durch Erfahren von intersubjektiver Anerkennung“ mitgedacht, z. B. Identitätsbildung.

3 Die lerntheoretische Argumentation wäre: „Aktiv an der Gestaltung des Lernens beteiligt sein ist eine wesentliche Voraussetzung für einen guten und sinnvollen Lernprozess, weil Lernen als aktiv aneignender Prozess verstanden wird“ (Bastian 2009, S. 7). Jene lerntheoretische Überlegungen werden im Rahmen der Arbeit zum Teil unter dem Ziel der „Qualifizierung“ aufgegriffen.

(vgl. Goffman 2002). Im Folgenden wird auf den Anerkennungsbegriff von Honneth zurückgegriffen (vgl. Honneth 1992). Dahinter steht der Gedanke, dass die Konkretisierung eines negativ besetzten Konzepts wie Diskriminierung bei der Beschreibung davon, wie Schulkultur idealerweise gestaltet sein sollte, nur begrenzt weiterhelfen würde (für die Verwendung des Honneth'schen Anerkennungsbegriff in Bezug auf Inklusion siehe Helsper et al. 2001; Sandring 2013). Im Folgenden wird daher argumentiert, dass das Erfahren von Anerkennung eine Gelingensbedingung für die Überwindung von Diskriminierung darstellt.

Bei Honneth wird Anerkennung „als ein Akt der Affirmation von positiven Eigenschaften menschlicher Subjekte“ (Balzer 2014, S. 578) verstanden. Demnach ist das Erfahren von Anerkennung für die Entwicklung von Selbstachtung und Selbstwertgefühl notwendig (Honneth 1992, S. 220; siehe auch Prengel 2006, S. 60). Dabei wird dem Anerkennungskonzept eine Doppelrolle zugewiesen:

„Einerseits scheint sie – als fehlende Anerkennung – der Grund gesellschaftlicher Ungleichheit und Marginalisierung zu sein, andererseits scheint sie zugleich – als Inbegriff der pädagogischen Strategie – auch der Weg der (Auf-)Lösung eben dieser Probleme zu sein“ (Balzer und Ricken 2010, S. 62).

In diesem Zusammenhang unterscheidet Honneth zwischen drei Anerkennungssphären: emotionaler Anerkennung; rechtlicher Anerkennung und der sozialen Wertschätzung (vgl. Honneth 1992; siehe auch Balzer 2014):

Emotionale Anerkennung: Darunter werden alle Beziehungen verstanden, die „aus starken Gefühlsbindungen zwischen wenigen Personen bestehen“ (Honneth 1992, S. 153). „Die Anerkennung [muss, ap] hier selber den Charakter affektiver Zustimmung und Ermutigung besitzen“ (ebd., S. 153).

Rechtliche Anerkennung: „Die Rechtssubjekte erkennen sich dadurch, dass sie dem gleichen Gesetz gehorchen, wechselseitig als Personen an, die in individueller Autonomie über moralische Normen vernünftig zu entscheiden vermögen“ (ebd., S. 177).

Soziale Wertschätzung: „Im Unterschied zur rechtlichen Anerkennung in ihrer modernen Gestalt, so haben wir gesehen, gilt die soziale Wertschätzung den besonderen Eigenschaften, durch die Menschen in ihren persönlichen Unterschieden charakterisiert sind“ (ebd., S. 197). Diese muss auf „intersubjektiv verbindliche Weise zum Ausdruck“ (ebd., S. 197) gebracht werden.

Es gibt diverse Interpretationen davon, wie Honneths Anerkennungssphären auf den pädagogischen Kontext übertragen werden können (z. B. Helsper et al. 2005; Lussi und S. Huber 2015).

Emotionale Anerkennung: Hier lässt sich nur schwer auf das Verhältnis zwischen Lehrkräften und Schüler_innen übertragen. Helsper verwendet das Bild einer

„prinzipiell positiv getönten Haltung des Interesses an der Person, den Lernwegen, Lernkrisen und Bildungsprozessen der Schüler, die auf einem an der ‚Sache‘, also der Ermöglichung umfassender Bildungsprozesse, orientierten Vertrauen basiert“ (Helsper et al. 2001, S. 32).

Diese Interpretation entfernt sich jedoch sehr stark von Honneths' Verständnis. So wäre die Haltung gegenüber Schüler_innen einseitig und es kann nicht von intersubjektiver Anerkennung gesprochen werden. Die begrenzte Anwendbarkeit der Sphäre bestätigt indirekt auch die empirischen Studien von Lussi und Huber, wonach sich Formen der emotionalen

Anerkennung⁴ im Rahmen von Peerbeziehungen beobachten lassen, aber kaum zwischen Lehrkräften und Schüler_innen (vgl. Lussi und S. Huber 2015).

Rechtliche Anerkennung: Hier geht es im pädagogischen Kontext um „gerechte Behandlung und der Zubilligung gleicher Rechte und Möglichkeiten“ (Helsper et al. 2001, S. 32).

„In welcher Form allen Schülern unabhängig von Geschlecht, Ethnie, Herkunft, Glaubensüberzeugungen etc. Möglichkeiten des gleichberechtigten Zuganges zu Lern- und Bildungsprozessen ermöglicht und gleiche Chancen der Partizipation eröffnet werden, ist für die konkrete Ausformung der Schulkultur zentral“ (ebd., S. 32).

Hier lässt sich Graumanns Interpretation von „rechtlicher Anerkennung“ ergänzen: So geht es in inklusiven Settings darum, dass Kinder erfahren, dass sie Rechte haben (Graumann 2012, S. 90). Hierbei betont sie das Recht auf Selbstbestimmung.

Soziale Wertschätzung: „Über diese Anerkennungsform der konkreten Person aufgrund ihrer spezifischen Leistungen, Eigenschaften, Haltungen, ihres Lebensstils und ihrer Selbstdarstellung konstituiert sich die Wertschätzung der Person in den Augen bedeutsamer anderer“ (Helsper et al. 2001, S. 33). Dabei basieren Formen der sozialen Wertschätzung nach Lussi und Huber im schulischen Kontext insbesondere auf „der Anerkennung individueller Leistungen und Fähigkeiten“ (Lussi und S. Huber 2015). Sandring macht in diesem Zusammenhang darauf aufmerksam, dass der Fokus auf Leistungen im Kontext von Bildungsungleichheiten kritisch zu sehen ist, da die unterschiedlichen Voraussetzungen der Schüler_innen nicht berücksichtigt werden. Graumann versucht diesen Aspekt zu berücksichtigen, indem sie „soziale Wertschätzung“ im inklusiven Setting folgendermaßen beschreibt: Es geht um die Überwindung einer Defizitorientierung sowie einer „wertschätzenden Einstellung gegenüber allen besonderen Fähigkeiten, Zugehörigkeiten und Lebensentwürfen“ (Graumann 2012, S. 92). Dies impliziert, dass Schüler_innen eine wertschätzende Haltung entgegengebracht wird, die unabhängig davon ist, welche schulischen Leistungen sie erbringen. Anknüpfend an Graumanns Konkretisierung der Anerkennungssphären für den schulischen Kontext bedeutet das Erfahren von intersubjektiver Anerkennung als ein Ziel von Inklusion, dass Schüler_innen sich ihrer Rechte bewusst sind und ihnen in der Interaktion mit anderen, z. B. mit Mitschüler_innen oder Lehrkräften, eine wertschätzende Haltung entgegengebracht wird.

Daran anknüpfend ist das Konzept der „egalitären Differenz“ nach Prengel (2001) zu nennen: „Im Gedanken der egalitären Differenz kommt Kritik an hierarchischen Strukturen auf der gesellschaftlichen wie auf der persönlichen Ebene zum Ausdruck“ (ebd., S. 94). Dabei beinhaltet das Konzept „die Entfaltung des Verschiedenen auf der Basis gleicher Rechte“ (Prengel 2006, S. 49). Demnach ist dem Konzept inhärent, dass alle Schüler_innen gleichwertig sind und zugleich in ihrer Unterschiedlichkeit anerkannt werden.

Der Honneth'sche Anerkennungs begriff ist jedoch abzugrenzen von anderen Verwendungen des Anerkennungs begriffs (für eine Übersicht siehe Balzer 2014). Nach Balzer und Ricken ermöglicht der positiv besetzte Anerkennungs begriff von Honneth zwar, dass „Erscheinungen der fehlenden Anerkennung oder Demütigungs- und Missachtungserfahrungen problematisierbar werden“ (Balzer und Ricken 2010, S. 54). Jedoch sehen Balzer und Ricken mit Verweis auf

⁴ Lussi und Huber bleiben hier nah an der Definition von Honneth und verstehen unter emotionaler Anerkennung „emotionale Zuwendungen oder Bindungen“ (Lussi und S. Huber 2015).

Emcke kritisch, dass Anerkennung in Form von kulturell codierten Zuschreibungen stattfindet, bestehende Machtstrukturen somit reproduziert werden (vgl. Balzer 2014, S. 578; siehe auch Balzer und Ricken 2010, S. 53). In diesem Fall wird Anerkennung als ein „spezifisches Strukturmoment einer jeden menschlichen Kommunikation und Praktik“ (ebd., S. 73) verstanden. Ein weiterer Kritikpunkt am Honneth'schen Anerkennungs-begriff ist nach Balzer und Ricken, dass dieser die Widersprüche im pädagogischen Handeln nicht sichtbar macht (vgl. ebd., S. 55). Demnach hat Schule die Aufgabe Schüler_innen zu fördern und zu unterstützen, aber auch zu sanktionieren und zu selektieren.

„Genau diese pädagogische Ambivalenz aber wird durch eine durchgängige positive Aufladung von Anerkennung mindestens übersprungen und mit Blick auf die Praxis erheblich beschwert, laufen doch regulierende, negierende wie sanktionierende Handlungen dann allzu leicht Gefahr, als Abwertungen und Missachtungen codiert zu werden“ (ebd., S. 55).

Dieser Kritikpunkt weist auf eine Grenze hin, die generell bei dem hier verwendeten Verständnis von Inklusion besteht: So liefert dieses keine Antworten darauf, inwiefern die durch Inklusion angestrebten Ziele erreicht werden sollen und zugleich eine weitere gesellschaftliche Funktion von Schule, die Allokationsfunktion, erfüllt werden soll (vgl. Moser 2017). Hier bleibt eine Leerstelle bestehen.

Trotz der Kritik am Honneth'schen Anerkennungs-begriff wird im Rahmen der Arbeit auf diesen zurückgegriffen. Gleichwohl werden die von Balzer und Ricken eingebrachten Kritikpunkte berücksichtigt, wenn es um die Operationalisierung von Differenzlinien in den Bewertungsinstrumenten der Schulinspektionen geht.

2.2.4 Inklusive Perspektive auf Menschenrechte

Bislang wurde dargestellt, was im Rahmen dieser Arbeit unter „Teilhabe“ und „Überwindung von Diskriminierung“ im schulischen Kontext verstanden wird: Qualifizierung, politische Partizipation sowie Erfahren von Anerkennung. Indem diese Konzepte im Rahmen dieser Arbeit Ziele von Inklusion darstellen, geht damit eine bestimmte Fokussierung einher: Die Situation von allen Schüler_innen wird in den Blick genommen.

Hierbei ist zu berücksichtigen, dass es in der Literatur unterschiedliche Positionen gibt, wie die Gesamtgruppe an Schüler_innen klassifiziert wird, auf welcher Ebene dies passiert und inwiefern bestehende Zuschreibungen kritisch hinterfragt werden. Beispielsweise gibt es Inklusionsverständnisse, die auf einer dichotomen Unterscheidung zwischen „Behinderung“ bzw. „Nicht-Behinderung“ basieren. Andere Autor_innen streben eine Überwindung der Zwei-Gruppen-Theorie an (vgl. Boban und Hinz 2003) (nähere Ausführung → Abschnitt 2.2.4).

Im Rahmen dieser Arbeit liegt der Fokus auf den individuellen Schüler_innen. Dies erlaubt es eine Vielzahl von Differenzlinien in den Blick zu nehmen, die in Bezug auf Teilhabe und Überwindung von Diskriminierung von Relevanz sind. Durch eine intersektionale Perspektive ist es darüber hinaus möglich nicht nur einzelne Differenzlinien in den Blick zu nehmen, sondern auch deren Verschränkung (vgl. Crenshaw 1989), z. B. Menschen, die den Rollstuhl zur Fortbewegung nutzen und sich zugleich als weiblich definieren. So ergeben sich aus der Verschränkung von Attributen andere Formen der Diskriminierung sowie andere Bedürfnisse in Bezug auf Teilhabe. Des Weiteren ist zwischen Fremd- und Selbstpositionierung zu unterscheiden (z. B. Emcke 2010): Es gibt kollektive Identitäten, die selbst gewählt sind, sowie solche, die durch Zuschreibungen von außen entstehen (vgl. ebd.). Beide sind im schulischen Kontext von Relevanz: So ist das Wahrnehmen und Wertschätzen von selbstgewählten Identitäten im Sinne

von intersubjektiver Anerkennung in Bezug auf die Identitätsbildung von Relevanz. Beispielsweise die Frage, wie sich einzelne Schüler_innen in Bezug auf Gender verorten und wie sie von anderen angesprochen werden möchten. Fremdzuschreibungen sind im schulischen Setting insbesondere von Relevanz, wenn diese mit stigmatisierenden Zuschreibungen einhergehen bzw. mit Selbsteinschätzungen nicht übereinstimmen. Im Rahmen dieser Arbeit wird darüber hinaus die Setzung vorgenommen, dass Differenzlinien sozial konstruiert sind: Sie entstehen in der Interaktion zwischen Individuum und gesellschaftlichen Strukturen (vgl. soziales Modell von Behinderung in der BRK).

Im Folgenden geht es nun darum, welche Implikationen eine inklusive Perspektive auf die oben genannten Ziele von Inklusion hat und was das für die Einzelschulebene bedeutet. Abschließend wird näher darauf eingegangen, was eine inklusive Perspektive in Bezug auf die Beschreibung von Lerngruppen und Benennung von Schülergruppen beinhaltet.

Ziel von Inklusion: Individuell bestmögliche Qualifizierung aller: Eine inklusive Perspektive auf das Konzept der Qualifizierung würde implizieren, dass alle Schüler_innen individuell bestmöglich qualifiziert werden. Wie lässt sich dieses Ziel in den Instrumenten der Schulinspektion operationalisieren? Nach Black-Hawkins ist es notwendig sich hierfür von folgendem Leistungsverständnis zu distanzieren:

„Standards in education are often considered to be a set of minimum performance criteria, usually in basic areas of the curriculum (...) that specify what all children should know, understand and be able to do by certain ages“ (Black-Hawkins et al. 2007, S. 21).

Von diesem Leistungsverständnis ausgehend hat Schule die Aufgabe, dass möglichst alle Schüler_innen zu einem bestimmten Zeitpunkt ein gewisses Leistungsniveau erreichen (hierzu kritisch ebd., S. 21). Jenes Verständnis spiegelt sich im Bildungsmonitoring beispielsweise in Lernstandserhebungen wie den IQB-Vergleichsarbeiten wider. Die Ergebnisse von Schulleistungstests lassen Aussagen darüber zu, inwiefern Schüler_innen über diejenigen Kompetenzen verfügen, die für notwendig erachtet werden, um im weiteren Lebensverlauf am gesellschaftlichen Leben teilhaben zu können. Darüber hinaus sind Querschnittsmessungen der Leistungen auch im Hinblick auf die Allokationsfunktion von Schule aussagekräftig.

Um den Aspekt der individuell bestmöglichen Qualifizierung zu berücksichtigen, misst die englische Schulinspektion (OFSTED)⁵, welche Fortschritte die einzelnen Schüler_innen im Laufe ihrer Schulzeit machen (vgl. Piezunka 2017). Dabei beziehen sie ein, über welche Kenntnisse Schüler_innen zu Beginn ihrer Schulzeit verfügen. Bei der Bewertung des Fortschritts orientieren sie sich an einem vorformulierten Soll-Wert⁶. Demnach gehen sie davon aus, dass Aussagen darüber möglich sind, ob der beobachtete Lernfortschritt von Schüler_innen „normal“ sei (hierzu kritisch Bradbury und Roberts-Holmes 2017). Dies widerspricht allerdings dem hier verwendeten Verständnis von Inklusion, weil dieses auf der Annahme basiert, dass es individuell variieren kann, was als bestmöglicher Fortschritt gesehen wird.

Durch Lernstandserhebungen wie den Vergleichsarbeiten sind demnach keine Aussagen darüber möglich, inwiefern die Schüler_innen eine individuell bestmögliche Entwicklung vollzogen haben. Darüber hinaus kann die Orientierung an Schulleistungstests zu nicht intendierten

5 Die Informationen zu England basieren auf einem Artikel, den die Autorin dieser Arbeit für den Sammelband „Inklusion und Schulentwicklung“ (hrsg. von Vera Moser und Marina Egger in 2017) geschrieben hat (Piezunka 2017).

6 Bei manchen Schülergruppen, z. B. Schüler_innen mit „special needs“ wird keine kriteriale, sondern eine soziale Bezugsnorm angewandt.

Wirkungen führen, die dem Inklusionsgedanken widersprechen: Mit Verweis auf die USA argumentiert Biewer, dass die „öffentliche Zurschaustellung ungünstiger Schülerleistungen einer Schule“ (Biewer 2012, S. 14) zur Folge haben kann, dass Eltern versuchen Schulen zu meiden, die in solchen Vergleichsarbeiten schlecht abschneiden. Dies führt dazu, dass bereits bestehende sozialräumliche Disparitäten wachsen (vgl. ebd., S. 14). Darüber hinaus kann es sein, dass manche Schulen als Reaktion auf die Lernstandserhebungen nur noch Schüler_innen aufnehmen, die für leistungsstark gehalten werden (vgl. Disability News 2017). Dahinter steht der Gedanke, dass diese im Vergleich zu anderen Schüler_innen leichter bestimmte Bildungsstandards erreichen. Des Weiteren beschreiben Jarke und Breiter (2016), dass sich Lehrkräfte im US-amerikanischen Bildungssystem darum bemühen, innerhalb einer Lerngruppe insbesondere diejenigen Schüler_innen zu fördern, die sich zwischen zwei Kompetenzstufen befinden. In Bezug auf das Erreichen einer höheren Kompetenzstufe ist bei diesen Schüler_innen im Vergleich zu ihren Mitschüler_innen weniger Aufwand notwendig. Auch dies stellt eine Reaktion auf die dort dominierenden Lernstandserhebungen dar. Abschließend lässt sich festhalten, dass Lernstandserhebungen im Querschnitt mit dem hier verwendeten Inklusionsverständnis kompatibel sind, aber sie können (nicht-)intendierte Wirkungen haben, die diesem widersprechen.

Nach Black-Hawkins et al. ist es daher notwendig in Bezug auf Inklusion von einem anderen Leistungsverständnis auszugehen (2007). So geht es darum, dass sich alle in Abhängigkeit von ihren Fähigkeiten und Bedürfnissen bestmöglich entwickeln (vgl. ebd., S. 21). Dies kann im Einzelfall bedeuten, dass einzelne Schüler_innen große Lernfortschritte machen, aber trotzdem die Bildungsstandards nicht erreichen, und dass andere Schüler_innen weitaus mehr lernen als in den Bildungsstandards vorgesehen. Aus der Inklusionsperspektive wäre demnach eine gute Leistungsentwicklung nicht gegeben, wenn im Durchschnitt möglichst viele Schüler_innen bestimmte Bildungsstandards erreichen, sondern wenn alle Schüler_innen sich bestmöglich entwickeln. In Bezug auf die Outputebene stellt sich jedoch die Frage, wie das Ziel „individuell bestmögliche Qualifizierung“ operationalisiert werden kann. So lässt sich nicht vorhersagen, was der bestmögliche Lernfortschritt für die einzelnen Schüler_innen wäre. So können Schulinspektionen lediglich auf der Prozessebene überprüfen, inwiefern „die Gelingensbedingungen für eine bestmögliche Qualifizierung aller gegeben sind, z. B. durch individuelle Lernentwicklungspläne im Unterricht“ (Piezunka 2017). Dabei schließen Formen der inneren Differenzierung nicht aus, dass kooperative Lernformen möglich sind, z. B. durch das Arbeiten an einem gemeinsamen Gegenstand (vgl. Feuser 1999). Im nächsten Schritt stellt sich die Frage, welche Qualifikationen Kinder und Jugendliche in der Schule erwerben sollen. Im Hinblick auf die Fragestellung dieser Arbeit ist hier keine Festlegung notwendig. So wurde bereits durch gesetzliche Vorgaben auf Bundeslandebene entschieden, welche Kompetenzen die Schüler_innen erwerben sollen. Demnach haben Schulinspektionen hierbei keinen Handlungsspielraum.

Das Ziel „individuell bestmögliche Qualifizierung“ wird berücksichtigt, wenn in Bezug auf Lernprozesse die Entwicklung des Individuums im Vordergrund steht. Dagegen ermöglichen Lernstandserhebungen keine Aussagen über individuell bestmögliche Qualifizierung bzw. solche Messungen können nicht-intendierte Wirkungen haben, die dem hier verwendeten Inklusionsverständnis widersprechen.

Ziel von Inklusion: Politische Partizipation aller Schüler_innen: Als weiteres Ziel von Inklusion wird im Rahmen dieser Arbeit die Möglichkeit der Mitbestimmung gesehen (siehe Deutsche UNESCO-Kommission 2014; vgl. Boban und Hinz 2003; J. Simon und T. Simon 2013; Platte 2014, S. 80; Black-Hawkins et al. 2007, 48f.). Durch die inklusive Perspektive

liegt der Fokus auf der Frage, wer an partizipativen Prozessen beteiligt ist (vgl. Reitz 2015). In diesem Zusammenhang sind zwei Aspekte von Relevanz: Alle Schüler_innen verfügen über das gleiche Stimmrecht. Dies birgt im konkreten Handeln die Gefahr, dass die Bedürfnisse von Minderheiten, die nicht nur aufgrund ihrer gesellschaftlichen Stellung, sondern auch quantitativ eine Minderheit darstellen, zu wenig gehört werden. Politische Partizipation aus Inklusionsperspektive impliziert daher auch, dass die Heterogenität der Schülerschaft in indirekten Verfahren repräsentiert ist und dass die Situation von Minderheiten besonders berücksichtigt wird. Dabei ist zu berücksichtigen, dass es in Abhängigkeit davon, worüber eine Entscheidung getroffen wird, variiert, welche kollektive Identität für relevant erachtet wird. Des Weiteren stellt sich auf der Prozessebene die Frage, inwiefern es für alle Schüler_innen die Möglichkeit gibt, ihr Recht auf Partizipation auszuüben (vgl. ebd., S. 7). Dies ist unabhängig davon, ob sie dies tatsächlich wahrnehmen. Abschließend ist noch zu erwähnen, dass die zweite Anerkennungssphäre nach Honneth, die rechtliche Anerkennung, bei dem Aspekt politischer Partizipation, insofern mitgedacht wird, als dass Partizipation auch die Anerkennung als autonomes und selbstbestimmtes Individuum impliziert (vgl. Moldenhauer 2015, S. 37).

Ziel von Inklusion: Erfahren von intersubjektiver Anerkennung: Im Folgenden steht die soziale Anerkennung, d. h. eine wertschätzende Haltung gegenüber den Schüler_innen, im Mittelpunkt. Dabei bedeutet dies für den schulischen Kontext, dass diese unabhängig davon ist, welche Leistungen erbracht werden. In Bezug auf Anerkennung besteht allerdings die Gefahr, dass eine positiv intendierte Haltung gegenüber bestimmten Merkmalen von Schüler_innen auch zur Reproduktion von Machtverhältnissen führen kann, z. B. stereotype Zuschreibungen (→ Abschnitt 2.2.4).

Inklusive Perspektive auf Schüler_innen als Querschnittsdimension: Aus dem hier verwendeten Verständnis von Inklusion ergibt sich – wie bereits dargestellt – eine bestimmte Perspektive auf die Gruppe an Schüler_innen: So ist die Situation aller Schüler_innen von Relevanz. Um diverse Differenzlinien sowie deren Verschränkung in den Blick nehmen zu können, werden die Schüler_innen in ihrer Individualität gesehen. Jene Perspektive positioniert sich kritisch gegenüber der linearen Aufzählung von einzelnen Differenzlinien, da dadurch die Verschränkung von Differenzlinien bei einzelnen Personen nicht in den Blick genommen werden kann und zum anderen werden bestimmte Etikettierungen aufgrund der stigmatisierenden Wirkung kritisch gesehen (vgl. Boban et al. 2014; siehe auch Moser 2012a; Prengel 2003). In Bezug auf die Operationalisierung von Inklusion in den Instrumenten der Schulinspektionen würde diese Perspektive implizieren, dass die Situation aller Schüler_innen in den Blick genommen wird und Gruppenzugehörigkeiten nicht mehr explizit benannt werden. Dies birgt jedoch die Gefahr, dass Differenzlinien, die in Bezug auf Teilhabe und Überwindung von Diskriminierung von Relevanz sind, unsichtbar werden. Dies hat Butler (2015) in Bezug auf die Verwendung der politischen Forderung „all lives matter“ (anstatt „black lives matter“) näher ausgeführt. So argumentiert sie, dass die Aussage „all lives matter“ nicht die bestehende Problematik erfassen würde:

„It is true that all lives matter, but it is equally true that not all lives are understood to matter which is precisely why it is most important to name the lives that have not mattered, and are struggling to matter in the way they deserve“ (ebd.).

Auf den schulischen Inklusionsbegriff übertragen, bedeutet dies, dass das Erreichen der Ziele von Inklusion für alle Schüler_innen angestrebt wird, jedoch gibt es bestimmte Personengruppen

bei denen das Erreichen dieser Ziele besonders in Gefahr ist. Demnach ist es notwendig diese Differenzlinien explizit zu benennen. So gelten für sie andere Bedingungen:

„But to make that universal formulation concrete, to make that into a living formulation, one that truly extends to all people, we have to foreground those lives that are not mattering now, to mark that exclusion, and militate against it“ (Yancy und Butler 2015).

Demnach ist es notwendig Gruppenzugehörigkeiten explizit zu benennen, um ihre Vulnerabilität sichtbar zu machen. Aus dieser Argumentation heraus liegt es nahe Differenzlinien explizit zu benennen. Darüber hinaus gibt es zwei weitere Aspekte, die für die explizite Benennung von Gruppenzugehörigkeiten sprechen: Wie bereits oben angedeutet, können Gruppenzugehörigkeiten im positiven Sinne identitätsstiftend sein. Beispielsweise gibt es Menschen, die zur Kommunikation primär die Deutsche Gebärdensprache verwenden und sich als Angehörige dieser Kultur verstehen. Die Zugehörigkeit zu dieser Gruppe ist in Bezug auf die eigene Identitätsbildung positiv besetzt und selbst gewählt. In diesem Fall impliziert die explizite Nennung von Differenzlinien keine Stigmatisierung, sondern macht Kulturen sichtbar. Darüber hinaus ermöglicht die explizite Verwendung von Differenzlinien bestehende Wissensbestände zu nutzen (vgl. Prengel 2003, S. 43). Beispielsweise kann auf bestehendes Wissen über die jeweilige Differenzlinie zurückgegriffen werden, um eine Person bestmöglich zu unterstützen. In diesem Zusammenhang argumentiert Prengel jedoch, dass der Rückgriff auf Differenzlinien lediglich eine erste Orientierung darstellen sollte (vgl. ebd., S. 34). Daraus lassen sich für den individuellen Einzelfall keine direkten Handlungsschritte ableiten (vgl. ebd., S. 34).

Die Kritikpunkte an der expliziten Verwendung von Differenzlinien bleiben aus der Inklusionsperspektive heraus bestehen: So ist stets kritisch zu hinterfragen, wann die jeweiligen Gruppenzugehörigkeiten von Relevanz sind. Daran anknüpfend stellt sich auch die Frage, welche Gruppenzugehörigkeiten im schulischen Kontext eine Rolle spielen und inwiefern damit pauschale Zuschreibungen einhergehen. Die Entscheidung darüber, welche Differenzlinien in Bezug auf Teilhabe und Diskriminierung zu berücksichtigen sind, kann temporal (z. B. Rudolf 2017, S. 31) und lokal variieren. In Bezug auf Deutschland zeigt sich beispielsweise, dass manche Differenzlinien im Laufe der Zeit aufgrund von soziodemographischen Veränderungen an Bedeutung gewonnen haben, z. B. Deutsch als Zweitsprache. Des Weiteren kann es auch variieren, wer benachteiligt wird: So wurde in den 60er Jahren im Kontext von Benachteiligung im Bildungssystem u. a. von Picht (1964) noch vom katholischen Arbeitermädchen vom Lande gesprochen, während heute der Arbeitersohn mit Migrationshintergrund aus der Großstadt (vgl. Klemm 2014) als Bildungsverlierer deklariert wird. Des Weiteren ist es unter anderem von Forschung und Politik abhängig, welche Differenzlinien bzw. Ausprägungen Aufmerksamkeit erfahren. Beispielsweise gibt es in Deutschland im Vergleich zum angloamerikanischen Bereich wenig Forschung zur Diskriminierung von Schüler_innen mit hohem Körpergewicht, d.h. nicht, dass es im Fall von Deutschland keine Diskriminierung gibt, sondern erstmal nur, dass wenig Wissen darüber vorhanden ist.

In Anlehnung an die Ziele von Inklusion sowie die gesellschaftlichen Funktionen von Schule bieten sich folgende Kriterien an, um herauszuarbeiten, ob Differenzlinien bzw. ein einzelnes Attribut von Schüler_innen in Bezug auf Inklusion relevant sind:

Kriterium 1: Haben die Attribute der Schüler_innen Einfluss auf den Lernprozess, z. B. Vorkenntnisse? (vgl. Piezunka 2015)

Kriterium 2: Gibt es stereotype Vorstellungen darüber, dass diese Einfluss auf den Lernprozess haben, z. B. Schüler_innen die nach der Definition des Statistischen Bundesamts „Migrationshintergrund“ aufweisen? (vgl. ebd.)

Kriterium 3: Gibt es stereotype Zuschreibungen, die die betroffenen Personen als verletzend empfinden, z. B. sexistische Äußerungen?

Kriterium 4: Gibt es Attribute, die nach der Meinung von vielen Menschen als von der gesellschaftlichen Norm abweichend angesehen werden bzw. sich dies in den bestehenden Strukturen widerspiegelt, z. B. Homosexualität (vgl. ebd.; in Bezug auf heteronormative Strukturen siehe Hartmann 2012).

Generell haben Messinstrumente das Potenzial unterschiedliche Formen von Diskriminierung, die auf sozial konstruierten Gruppenzugehörigkeiten basieren, sichtbar zu machen. Auf der anderen Seite besteht die Gefahr, dass durch Messinstrumente, die auf Gruppenzugehörigkeiten basieren, diskriminierende Klassifikationen reproduziert werden (vgl. Supik 2017).

2.2.5 Abgrenzung zu anderen Verständnissen

Das hier verwendete Verständnis von Inklusion lässt sich von anderen Verständnissen von Inklusion abgrenzen, ...

- ...die sich **nur auf Schüler_innen mit zugeschriebener Behinderung** bzw. diagnostiziertem sonderpädagogischem Förderbedarf beziehen. Dies würde insofern zu kurz greifen, als dass es auch in Bezug auf andere Differenzlinien empirische Nachweise gibt, dass deren Teilhabe in Gefahr ist. Wie bereits oben erwähnt, ist der Fokus auf zugeschriebene Behinderung dann legitim, wenn es um die Implementation der BRK geht, da nur die Rechte dieser Gruppe durch die BRK geschützt werden (vgl. Piezunka et al. 2018).
- ...wodurch die **Unterscheidung zwischen zwei Gruppen** – z. B. Schüler_innen mit und ohne zugeschriebener Behinderung – aufrechterhalten wird. Dies ist aus mehreren Gründen kritisch zu sehen: Beispielsweise ist die Gruppe an Schüler_innen mit zugeschriebener Behinderung in sich sehr heterogen (vgl. Boban und Hinz 2011) und intersektionale Verschränkungen können durch diese Perspektive nicht ausreichend in den Blick genommen werden. Demnach sind Aussagen, die auf dichotome Gruppeneinteilungen basieren, wenig aussagekräftig.
- ...**die sich auf die Lernentwicklung von Schüler_innen beschränken**. Ausgehend von den Funktionen nach Fend hat Schule nicht nur die Aufgabe Schüler_innen bei dem Erwerb von kognitiven Kompetenzen zu begleiten bzw. anzuleiten, sondern erfüllt innerhalb der Gesellschaft weitere Funktionen, z. B. Ausbildung der Urteilsfähigkeit.
- **..die in der Praxis nicht umsetzbar sind**, z. B. Inklusion als Utopie (vgl. Piezunka et al. 2017) bzw. Inklusion als Nordstern (vgl. Hinz 2010). Mit dem Fokus auf Schulinspektionen geht die Setzung einher, dass die Umsetzung von Maßnahmen in Bezug auf Inklusion für Einzelschulen realisierbar sind.

Das hier verwendete Verständnis von Inklusion weist jedoch Grenzen auf: Indem nur die Einzelschulebene betrachtet wird, können andere Ebenen des Bildungssystems nicht berücksichtigt werden. So kann es sein, dass auf höherer Ebene bestimmte Anforderungen an ein inklusives Bildungssystem nicht erfüllt werden, beispielsweise die wohnortnahe Beschulung aller Kinder

und Jugendlichen (vgl. Deutsches Institut für Menschenrechte 2014, S. 35; Deutsches Institut für Menschenrechte 2011, S. 2; siehe auch Avenarius 2012), oder die Nicht-Etikettierung von Schüler_innen (vgl. Biewer 2009, S. 193; Moser 2012a; Hinz 2007). Des Weiteren bleibt in Bezug auf die Allokationsfunktion von Schule (vgl. Fend 2006) eine Leerstelle. So hat Schule die Aufgabe zu selektieren bzw. soziale Positionen zuzuweisen, z. B. durch die Verteilung auf unterschiedliche Schultypen. Dies ist nach Moser (2017) mit der Umsetzung von Inklusion kaum vereinbar, da „ein positives Selbstkonzept sowie eine förderliche Lernatmosphäre unter dem Druck von Selektion und Allokation nur schwerlich herzustellen sind“ (ebd.). In diesem Zusammenhang stellt sich auch die Frage, inwiefern verschiedene Formen der Selektion, z. B. Leistungsbewertung, Verteilung auf unterschiedliche Schulformen, Wiederholung von Klassen, überhaupt notwendig sind, damit Schule diese Funktion erfüllen kann (z. B. Breidenstein 2017).

Wie bereits oben erwähnt, stellen die hier dargestellten Konzepte innerhalb der Literatur keine Neuentwicklung dar, sondern knüpfen an bereits bestehende Verständnisse von Inklusion an. So kann sowohl an Arbeiten vor und nach der Ratifizierung der BRK angeknüpft werden, z. B. Prengel (2006); Ainscow et al. (2006); Biewer (2009); Feuser (1995) und Hinz (2010). Des Weiteren zeigen sich Überschneidungen zu einem der vier Verständnisse von Inklusion, welches Piezunka et al. (2017) bei der Befragung von Inklusionsforscher_innen im deutschsprachigen Raum herausgearbeitet haben: So werden bei dem dritten Inklusionsverständnis mit der Bezeichnung <Anerkennung/Teilhabe/Wohlfühlen> die Schüler_innen in ihrer Individualität gesehen und neben dem Kompetenzerwerb werden in Bezug auf Inklusion als weitere Ziele soziale Integration und Anerkennung genannt (vgl. ebd.).

2.3 Forschungsstand: Umsetzung von Inklusion in Deutschland

Im Folgenden wird näher dargestellt, inwiefern Inklusion als politische Vorgabe in den einzelnen Bundesländern auf Einzelschulebene bereits umgesetzt ist. Dies ist insofern für die Arbeit relevant, als dass der situative Kontext die Handlungsoptionen von Schulinspektionen beeinflussen kann (nähere Ausführung → Kapitel 3).

Hierbei ist jedoch zu berücksichtigen, dass die Frage nach dem Stand der Umsetzung nur begrenzt zu beantworten ist: So ist hierfür **erstens ein Set an empirisch überprüfbaren Indikatoren notwendig**, die Inklusion als politische Vorgabe in Gänge und Komplexität abbildet (vgl. Piezunka et al. 2018). **Zweitens sind zusätzlich Benchmarks notwendig**, die in Bezug auf die einzelnen Indikatoren festlegen, wann ein Teilaspekt von Inklusion umgesetzt bzw. nicht umgesetzt ist. Um ein Beispiel zu nennen: Die European Agency for Development in Special Needs Education schlägt als Benchmark vor, dass eine Lerngruppe 80% des Unterrichts gemeinsam verbringen sollte (vgl. European Agency for Development in Special Needs Education 2011, S. 49). **Drittens bedarf es zur Beantwortung der Frage repräsentative Daten**, die Aussagen über die Umsetzung von Inklusion in einzelnen Bundesländer zulassen. So müssen bei der Erhebung auch Spezifika der einzelnen Bundesländer, z. B. Unterschiede in der statistischen Erhebung, berücksichtigt werden (vgl. Gresch und Piezunka 2015).

Die aktuelle Datenlage wird diesen Anforderungen nicht gerecht: Dies liegt unter anderem daran, dass es an einem allgemein anerkannten Set an empirisch überprüfbaren Indikatoren und dazugehörigen Benchmarks fehlt, um die Umsetzung von Inklusion als politische Vorgabe zu

messen. Erste Hinweise auf relevante Indikatoren ergeben sich aus der Staatenberichterstattung des Deutschen Instituts für Menschenrechte (DIMR) in Bezug auf die BRK (z. B. Deutsches Institut für Menschenrechte 2015). Hierbei ist jedoch zu berücksichtigen, dass sich das DIMR in seinen Publikationen auf die Gruppe an Menschen mit diagnostizierter Behinderung konzentriert, weil der Fokus ihrer Arbeit in diesem Fall auf der Umsetzung der BRK liegt. Damit würde nur ein Teilaspekt des in dieser Arbeit verwendeten Inklusionsverständnisses (→ Abschnitt 2.2) abgebildet werden. Darüber hinaus gibt es in Bezug auf Art. 24 BRK von Seiten des DIMR nur einzelne Indikatoren und es fehlt auch hier an einem Set an empirisch überprüfbareren Indikatoren, welches alle in Art. 24 BRK angesprochenen Aspekte in Bezug auf Deutschland in Gänze abdeckt (vgl. Piezunka et al. 2018). Des Weiteren ist festzustellen, dass es in Bezug auf den Stand der Umsetzung zwar einzelne Datensätze gibt, die Bundesländervergleiche zulassen, z. B. der Anteil an Schüler_innen, die exklusiv beschult werden (vgl. Bertelsmann-Stiftung 2014). Diese stehen jedoch zum Teil in der Kritik, weil sie bundesländerspezifische Unterschiede nur begrenzt berücksichtigen (z. B. Piezunka et al. 2016). Demnach lässt sich festhalten, dass die Frage nach dem Stand der Umsetzung von Inklusion nur begrenzt beantwortet werden kann, da es an einem allgemein anerkannten Set an empirisch überprüfbareren Indikatoren sowie darauf basierenden Datensätzen zu Deutschland fehlt.

Gleichwohl geben einzelne empirische Studien erste Anhaltspunkte dafür, wie es um die Umsetzung von Inklusion im Fall von Deutschland steht. So gibt es zwar eine Vielzahl von Studien aus den letzten Jahren, z. B. aus der soziologischen Ungleichheitsforschung, rechtliche Analysen sowie erziehungswissenschaftliche Arbeiten. Diese setzen sich jedoch nur mit Teilaspekten des in der Arbeit verwendeten Verständnisses von Inklusion auseinander, z. B. Kompetenzentwicklung, Unterrichtsgestaltung, soziale Integration. Aufgrund der hohen Anzahl an Studien würde es den Rahmen der Arbeit sprengen, den Forschungsstand hierzu vollständig aufzuarbeiten. Darüber hinaus wäre für die Entscheidung darüber, welche Forschung von Relevanz ist, eine Festlegung davon notwendig, wie die einzelnen Konzepte des oben beschriebenen Inklusionsverständnisses (→ Abschnitt 2.2) in empirischen Studien operationalisiert werden sollen, wie etwa das Konzept individuell bestmögliche Qualifizierung. Hierfür fehlt es an Festlegungen von bildungspolitischer Seite.

Im Folgenden wird daher beispielhaft auf einzelne Studien zurückgegriffen, um die Situation in Deutschland zu illustrieren. Dabei basierte die Auswahl der Studien, die im Folgenden beschrieben werden, erstens auf dem Kriterium, inwiefern sie die in Abschnitt 2.2 dargestellten Ziele von Inklusion (individuell bestmögliche Qualifizierung; Erfahren von intersubjektiver Anerkennung sowie politische Partizipation) abbilden und zweitens Aussagen darüber zulassen, inwiefern es Unterschiede zwischen verschiedenen Schülergruppen gibt. Drittens liegt der Fokus – soweit vorhanden – auf Studien, die bundesländerübergreifend Aussagen zulassen.

2.3.1 Individuell bestmögliche Qualifizierung

Die empirische Erfassung des Ziels „individuell bestmögliche Qualifizierung“ ist insofern herausforderungsvoll, als dass es hierfür einer Festlegung davon bedarf, was bestmögliche Qualifizierung im Einzelfall bedeutet (nähere Ausführung → Abschnitt 2.2.4). Als Messung von Lernprozessen bieten sich Kompetenzmessungen an: Diese geben zwar keine Informationen darüber, inwiefern auf individueller Ebene Schüler_innen bestmöglich qualifiziert wurden. Sie können jedoch nachweisen, inwiefern im Bildungssystem Disparitäten in Bezug auf bestimmte Differenzlinien existieren (z. B. Solga und Dombrowski 2009; Solga und R. Becker 2012).

So konnte insbesondere durch die PISA-Studie (z. B. Stanat et al. 2002) gezeigt werden, dass im Hinblick auf Kompetenzen verschiedene Differenzlinien eine Rolle spielen: In Bezug auf die Differenzlinie „soziale Herkunft“ zeigte sich beispielsweise im Bildungsbericht von 2016, dass Leistungsunterschiede von der sozialen Herkunft der Schülerinnen und Schüler abhängen (vgl. Autorengruppe Bildungsberichterstattung 2016, S. 92). Neben der sozialen Herkunft spielen noch weitere Differenzlinien in Bezug auf die Lernprozesse von Schüler_innen eine Rolle, z. B. Gender (für einen Überblick siehe Hadjar 2011), Behinderung (z. B. Powell 2007) oder Migrationshintergrund (für einen Überblick siehe Diehl et al. 2016). Darüber hinaus konnte durch Studien wie PISA-E nachgewiesen werden, dass es zwischen den Bundesländern Unterschiede darin gibt, wie eng der Zusammenhang zwischen Lernprozess und bestimmten Differenzlinien ist: So zeigte sich beispielsweise, dass in manchen Ländern wie Brandenburg oder Bayern geringere Zusammenhänge zwischen sozialer Herkunft und Kompetenzniveau bestehen als in anderen Bundesländern wie Baden-Württemberg und Bremen (vgl. Prenzel et al. 2005, S. 31).

Der Zusammenhang zwischen sozialer Herkunft und Kompetenzentwicklung wird unter anderem damit erklärt, dass Schüler_innen aus bildungsfernen Haushalten bei schulischen Herausforderungen weniger Unterstützung durch ihre Familie erhalten. Demnach gelingt es dem Bildungssystem nicht bestehende Ungleichheiten zu kompensieren. Es gibt jedoch auch Studien, die nachweisen, dass in schulischen Settings Ungleichheiten reproduziert bzw. geschaffen werden: Dies ist beispielsweise der Fall, wenn im Rahmen von experimentellen Designs nachgewiesen wird, dass Schüler_innen trotz gleicher Leistungen aufgrund eines Attributs bei der Notenvergabe benachteiligt werden (vgl. Sprietsma 2009).

Neben der Messung von Kompetenzen kann das Ziel „individuell bestmögliche Qualifizierung“ auch empirisch erfasst werden, indem Ergebnisse von institutionellen Selektionsprozessen näher betrachtet werden (für eine Übersicht siehe Die Antidiskriminierungsstelle des Bundes 2013). Ein möglicher Indikator für Selektionsprozesse ist der Anteil an Schüler_innen, die in exkludierenden Settings beschult werden: Beispielsweise ist dies in Sachsen-Anhalt für 9,2% und in Rheinland-Pfalz für 5,6% aller Schüler_innen der Fall (vgl. Autorengruppe Bildungsberichterstattung 2016, S. 81). Darüber hinaus arbeiten beispielsweise Radtke und Gomolla (2009) auf, inwiefern bestimmte administrative Regelungen oder etablierte Praktiken dazu führen, dass bestimmte Schülergruppen benachteiligt werden: Hierfür beschreiben sie den Fall, dass Schüler_innen aufgrund fehlender Deutschkenntnisse später eingeschult werden, ohne dass dabei sichergestellt wird, dass sie im vorschulischen Bereich zusätzliche Sprachförderung erhalten. Demnach spielen auch sprachliche Kenntnisse bei Selektionsprozessen eine Rolle.

Die ausgewählten Studien machen exemplarisch deutlich, dass im deutschen Bildungssystem in Bezug auf Lernprozesse Ungleichheiten bestehen, die mit einzelnen bzw. miteinander verschränkten Differenzlinien zusammenhängen. Dies gibt erste Hinweise darauf, dass eine individuell bestmögliche Qualifizierung im Sinne von Inklusion im deutschen Bildungssystem nicht gewährleistet ist. Darüber hinaus lässt sich durch bundesländerübergreifende Studien wie PISA-E nachweisen, dass sich die einzelnen Bundesländer darin unterscheiden, welche Bedeutung bestimmte Differenzlinien in Bezug auf den Lernprozess haben.

2.3.2 Erfahren von intersubjektiver Anerkennung

Als weiteres Ziel von Inklusion wird im Rahmen der Arbeit das „Erfahren von intersubjektiver Anerkennung“ aufgeführt. Hierfür können unter anderem solche Studien Informationen lie-

fern, die soziale Interaktionen im schulischen Kontext näher untersuchen. Jedoch gibt es nach Jennessen et al. (2013) nur „wenige empirische und belastbare Aussagen zum Diskriminierungsverhalten von Lehrkräften, Mitschüler_innen und auch Eltern sowie anderen am Schulleben Beteiligten“ (ebd., S. 56).

Als ein Beispiel ist die Studie von Streib-Brzić und Quadflieg (2011) zu nennen. Die beiden Autor_innen haben untersucht, welche Erfahrungen Kinder aus Regenbogenfamilien im schulischen Setting gemacht haben. Danach geben die meisten der von ihnen befragten Schüler_innen (N = 22) an, dass im schulischen Kontext „offen ausgetragene Konflikte, auch solche mit physischer Gewaltanwendung, eher alltäglich seien“ (ebd., S. 21). Demnach ist das Erfahren von intersubjektiver Anerkennung im Sinne der sozialen Wertschätzung in Bezug auf diese Gruppe von Schüler_innen nicht gegeben. Während die Studie von Streib-Brzić und Quadflieg die Situation einer spezifischen Gruppe näher untersucht, geht es im Rahmen der INTAKT-Studie allgemein um soziale Interaktionen zwischen Lehrkraft und Schüler_innen. Hierbei wurde festgestellt, dass ca. ein Fünftel der beobachteten Interaktionen in pädagogischen Settings verletzend sind (vgl. Prengel 2013). Weitere Studien, die mögliche Daten bieten, um Aussagen über das Ziel „Erfahren von intersubjektiver Anerkennung“ machen zu können, analysieren Normalitätsvorstellungen in Schulbüchern (z. B. zu Gender siehe Bittner 2011; zu Migration und Integration siehe Die Beauftragte der Bundesregierung für Migration, Flüchtlinge und Integration 2015).

Abschließend lässt sich festhalten, dass es eine Vielzahl von Studien gibt, die unterschiedliche Aspekte in Bezug auf das Erfahren von intersubjektiver Anerkennung empirisch untersuchen, z. B. Normalitätsvorstellungen oder Interaktionen zwischen diversen Akteuren im schulischen Setting. Bislang gibt es jedoch nach dem Wissensstand der Autorin dieser Arbeit keine Studie, die bundesländerübergreifend repräsentative Aussagen über die Qualität von sozialen Interaktionen in schulischen Settings machen kann.

2.3.3 Politische Partizipation

Um Aussagen darüber treffen zu können, inwiefern alle Kinder und Jugendliche die Möglichkeit haben im schulischen Setting mitzubestimmen, bietet es sich an, auf Studien zurückzugreifen, die Schüler_innen nach ihren subjektiven Einschätzungen zu Partizipationsmöglichkeiten befragen (für einen Überblick siehe Kurth-Buchholz 2011). Bei der Betrachtung des aktuellen Forschungsstands zeigt sich, dass eine Vielzahl von Studien erhebt, inwiefern Jugendliche sich außerhalb des schulischen Bereichs politisch engagieren, z. B. die Shell-Jugendstudie oder der DJI-Jugendsurvey. Im Rahmen der Arbeit geht es jedoch um Partizipationsmöglichkeiten in der Schule und um die Frage, inwiefern die Partizipationsmöglichkeiten zwischen einzelnen Schülergruppen variieren.

Hierzu hat das Deutsche Kinderhilfswerk im Jahre 2013 eine Studie zu Mitbestimmungsmöglichkeiten von Kindern und Jugendlichen durchgeführt. In Bezug auf den schulischen Bereich haben sie herausgefunden, dass ca. 55% der befragten Personen (N= 830) angeben, dass sie gar nicht mitbestimmen können (vgl. Deutsches Kinderhilfswerk e.V. 2013, S. 23). Hierbei schlüsseln sie ihre Ergebnisse nach Mädchen und Jungen sowie verschiedenen Altersgruppen auf, jedoch zeigen sich hierbei keine großen Unterschiede zwischen den einzelnen Gruppen.

Des Weiteren haben Grundmann und Kramer empirisch erhoben, wie Schüler_innen ihre Mitbestimmungsmöglichkeiten einschätzen. Danach können sie nachweisen, dass im Fall von Sachsen-Anhalt für fast alle befragten Schüler_innen Partizipationsmöglichkeiten im außerunter-

richtlichen Schulleben gewährleistet sind (vgl. Grundmann und Kramer 2001, S. 61; siehe auch Grundmann et al. 2000). Hierbei ist jedoch zu beobachten, dass die Partizipationsmöglichkeiten in Abhängigkeit von der Partizipationsform, z. B. Gestaltung der Schule oder Mitgestaltung der Hausordnung, variieren (vgl. ebd., S. 262). Anknüpfend an die Arbeiten von Grundmann et al. hat Böhm-Kasper (2006) Daten zu Sachsen-Anhalt und Nordrhein-Westfalen ausgewertet. Böhm-Kasper konnte zeigen, dass ein Zusammenhang zwischen der politischen Sozialisation im Elternhaus sowie der tatsächlichen Beteiligung in der Schule besteht (ebd.). Darüber hinaus konnte nachgewiesen werden, dass es auch signifikante Unterschiede in Bezug auf das Geschlecht gibt (ebd., 16f.).

Des Weiteren ist noch das Kinderpanel des Deutschen Jugendinstituts zu nennen, das vielfältige Partizipationserfahrungen im schulischen Alltag erfasst. Dabei sind auch Auswertungen auf der Bundesländerebene möglich: Hierbei zeigt sich, dass es zwischen den Bundesländern zwar keine signifikanten Unterschiede gibt, jedoch lässt sich als Tendenz herausarbeiten, dass „in den Stadt-Staaten Bremen und Hamburg die Mitbestimmung am größten ist. Auf Berlin trifft dies nicht zu.“ (Bacher et al. 2007). Des Weiteren arbeiten sie einen Zusammenhang zwischen Partizipationsmöglichkeiten in der Familie sowie in der Schule heraus: Wenn Möglichkeiten der Partizipation im Elternhaus gegeben sind, wirkt sich dies auch auf den schulischen Bereich aus (vgl. ebd.). Nichtsdestotrotz lässt sich im Rahmen der Studie kein signifikanter Zusammenhang zwischen schulischer Partizipation und der sozioökonomischen Herkunft von Schüler_innen herausarbeiten (vgl. ebd.).

Abschließend lässt sich festhalten, dass die aufgezeigten Studien in Bezug auf politische Partizipation im schulischen Settings zum Teil Unterschiede zwischen verschiedenen Schülergruppen deutlich machen, z. B. hinsichtlich des Elternhauses, Alters oder Geschlechts. Jedoch fehlt es in diesem Bereich an Studien, die noch weitere Differenzlinien, wie Behinderung, in den Blick nehmen.

2.3.4 Zwischenfazit

Die Frage, inwiefern Inklusion als politische Vorgabe in Deutschland umgesetzt ist, lässt sich – wie bereits oben dargestellt – nur begrenzt beantworten. Es lässt sich jedoch festhalten, dass in Bezug auf die oben genannten Ziele Unterschiede zwischen einzelnen Schülergruppen existieren. Abschließend stellt sich die Frage, welche Relevanz diese Erkenntnisse in Bezug auf diese Arbeit haben: **Erstens** lässt sich aus jenen Forschungsergebnissen begründen, dass auf Einzelschulebene Handlungsbedarf in Bezug Inklusion besteht und somit kann auch deren Berücksichtigung in den Instrumenten der Schulinspektionen legitimiert werden. So handelt es sich bei der Umsetzung von Inklusion nicht etwa um eine Norm, die in internationalen Menschenrechtskonventionen aufgeführt wird, aber im Fall von Deutschland bereits umgesetzt ist und daher nicht von Relevanz ist. **Zweitens** zeigte sich, dass sich die Bundesländer in Bezug auf einzelne Aspekte, z. B. Beschulung in exklusiven Settings sowie Kompetenzentwicklung, voneinander unterscheiden. Es fehlt jedoch an bundesländerübergreifenden Daten, um Aussagen über den Grad der politischen Partizipation und dem Erfahren von intersubjektiver Anerkennung machen zu können. In Bezug auf bestehende Unterschiede zwischen den Bundesländern stellt sich die Frage, ob sich die Unterschiede auch in den bundesländerspezifischen Instrumenten der Schulinspektionen widerspiegeln (→ Kapitel 7). **Drittens** hat dieses Unterkapitel auf der Metaebene deutlich gemacht, dass es unterschiedliche Formen davon gibt, wie Inklusion in der empirischen Bildungsforschung operationalisiert werden kann, da es aktuell

an einem allgemein anerkannten Set an empirisch überprüfbaren Indikatoren zur Umsetzung von Inklusion fehlt. Diese fehlende Orientierung an einem allgemein anerkannten Set stellt nicht nur eine Herausforderung für die empirische Inklusionsforschung dar, sondern auch für die Schulinspektionen.

3 Schulinspektionen in Deutschland

Im Rahmen dieses Kapitels wird dargestellt, wie es zur Einführung der aktuell existierenden Schulinspektionen gekommen ist, wie diese im Bildungssystem strukturell verankert sind und welche Funktionen ihnen zugeschrieben werden. Dies gibt Hinweise darauf, welche Rahmenbedingungen möglicherweise die Entwicklung von Messinstrumenten zu Inklusion prägen. Am Ende des Kapitels wird ein Überblick darüber gegeben, welche wissenschaftlichen Studien bereits zu Schulinspektionen existieren. Damit wird deutlich, inwiefern diese Arbeit an bestehende Studien anknüpfen kann bzw. Forschungslücken schließt. Zunächst wird jedoch der Ablauf eines Inspektionsverfahrens beschrieben. Dies ermöglicht einen Einblick in die Kernaufgabe von Schulinspektionen und macht sichtbar, welche Rolle die Messinstrumente hierbei spielen.

3.1 Beschreibung eines Inspektionsverfahrens

Im Folgenden werden die Schritte einer Inspektion näher beschrieben. Da sich die Inspektionsverfahren in den jeweiligen Bundesländern in einzelnen Punkten unterscheiden, z. B. im Ablauf, weicht der konkrete Fall auf Bundeslandebene zum Teil von dieser allgemein gehaltenen Beschreibung ab.

Einige Wochen vor einer Inspektion wird die Schule darüber informiert, wann diese stattfindet. In einigen Fällen gibt es auch ein persönliches Vorabgespräch, um den Schulen einen Überblick über den geplanten Schulbesuch zu geben (z. B. in Berlin). Parallel dazu tragen die Mitarbeiter_innen der Schulinspektion verschiedene Dokumente und Daten (z. B. Schulprogramm, Websiteauftritt, Daten zur Schülerschaft usw.) zusammen, um diese auszuwerten. Des Weiteren findet in vielen (z. B. Hamburg und Brandenburg), aber nicht in allen Bundesländern vor der Inspektion eine Onlinebefragung von Eltern, Lehrkräften und Schüler_innen statt. Diese hat zum Ziel möglichst repräsentative Daten über die Schule zu erheben und dient der Vorbereitung des Schulbesuchs: So vermitteln diese Daten bereits einen ersten Eindruck und ermöglichen die Bildung von Hypothesen.

Beim Schulbesuch gehen Mitarbeiter_innen der Schulinspektion für ca. zwei bis drei Tage an eine Schule, im Falle einer größeren Schule auch für einen längeren Zeitraum. Dabei besuchen sie für jeweils zwanzig Minuten (in Bayern: 40 Minuten) einen Großteil der Unterrichtsstunden, die an diesen Tagen stattfinden. Des Weiteren führen sie Gespräche mit verschiedenen schulischen Vertretern, wie Schüler_innen, Lehrkräften, Eltern, Hausmeister_innen, Erzieher_innen, Sozialarbeiter_innen sowie Assistent_innen. Bei ihrem Vorgehen orientieren sie sich an empirischen Verfahren der Sozialforschung und greifen beispielsweise auf leitfadengestützte Interviews zurück, um Daten zu erheben. Grundlage für die Evaluation ist dabei häufig ein Qualitätstableau, welches verschiedene Merkmale von guter Schule in Form von Indikatoren enthält (nähere Ausführung → Kapitel 6). Dieses Qualitätstableau bzw. die dazugehörigen Messinstrumente sind primär von den Mitarbeiter_innen der jeweiligen Schulinspektionen konzipiert worden. Dabei repräsentieren die jeweiligen Qualitätstableaus sowie die dazugehörigen

Messinstrumente bestimmte Verständnisse von Schulqualität und dienen den Schulen als Orientierung in Bezug auf die Frage, welche Anforderungen von Seiten der Bildungsadministration an sie formuliert werden.

Nach der Durchführung der Inspektion wird ein Bericht geschrieben, der sich am Qualitätstableau orientiert. Er beinhaltet in den meisten Fällen nicht nur Fließtext, sondern auch eine Auswertung der quantitativ erhobenen Daten. Auch hier gibt es Unterschiede zwischen den Bundesländern: So variiert es, wie die Ergebnisse an Schulen, Schulaufsicht und interessierte Öffentlichkeit kommuniziert werden. Gemein ist allen Bundesländern, dass die Schulen eine schriftliche Fassung des Inspektionsberichts erhalten. Häufig wird dieser auch mündlich der Schulgemeinschaft präsentiert. Zusätzlich wird in manchen Bundesländern (BE, BB, HH) noch eine gekürzte Version des Inspektionsberichts online veröffentlicht. Darüber hinaus unterscheiden sich die Bundesländer darin, wie die Schulaufsicht in den Prozess der Rückmeldung einbezogen wird und inwiefern die Schulen eine Rückmeldung zum Bericht geben können. Beispielsweise haben die Schulen in Berlin die Möglichkeit den online einsehbaren Inspektionsbericht zu kommentieren. Dieser Kommentar wird ebenfalls auf der Website der Senatsverwaltung veröffentlicht.

Nach Abschluss der Inspektion werden die meisten Schulen nach fünf bis sechs Jahren erneut evaluiert; so zumindest das erklärte Ziel der Bildungsadministrationen. Des Weiteren gibt es einzelne Bundesländer, in denen die Schulen aufgrund schlechter Ergebnisse bereits nach einem kürzeren Zeitpunkt nochmals inspiziert werden. Darüber hinaus wird in vielen Bundesländern von den Schulen erwartet, dass sie nach der Inspektion das Gespräch mit der Schulaufsicht suchen, um über mögliche Formen der Unterstützung und nächste Schritte zu sprechen. Abgesehen von dem kürzeren Inspektionsintervall und den Zielvereinbarungen mit der Schulaufsicht haben die Schulen in Deutschland – im Vergleich zu England oder den Niederlanden – keine größeren Sanktionen wie Schulschließung oder Personalwechsel zu befürchten. Aufgrund der niedrig angesetzten Sanktionen wird Deutschland daher im internationalen Vergleich als moderat bzw. als ein „low-stake“ – Verfahren beschrieben (vgl. Heinrich und Lambrecht 2016; siehe auch Bezem und Jacobsen 2018).

Abschließend lässt sich in Bezug auf die Messinstrumente festhalten, dass diese im Fall von Deutschland (Ausnahmen: MV und HB)¹ als wichtige Grundlage dienen, damit die Inspektor_innen die jeweiligen Schulen evaluieren können. Dies ist abzugrenzen von anderen Inspektionsmodellen, z. B. in Argentinien (vgl. ebd.), wo der persönliche Eindruck bzw. das subjektive Verständnis von guter Schule der jeweiligen Inspektor_innen entscheidend ist.

3.2 Einführung von Schulinspektionen

Bereits im 16. Jahrhundert etablierten sich die ersten kirchlichen Schulinspektionen auf Gebieten, die heute zu Deutschland gehören (vgl. Brüggemann et al. 2011). Danach kam es immer wieder zur Etablierung von Schulinspektionen, z. B. während der Weimarer Republik oder im Nationalsozialismus (vgl. ebd.). Brüggemann et al. (2011) argumentieren jedoch, dass sich die Funktion und das Format der Schulinspektionen im Laufe der Zeit stark gewandelt

¹ Im Rahmen dieser Arbeit werden nur die Bundesländer betrachtet, die mit standardisierten Instrumenten arbeiten (nähere Ausführung → Kapitel 6). Daher sind Mecklenburg-Vorpommern und Bremen nicht Teil der Stichprobe.

haben. So haben frühere Schulinspektionen den Fokus daraufgelegt, inwiefern von Seiten der Schule Gesetzesvorschriften eingehalten werden (vgl. Landwehr 2011, S. 55). Dagegen spielten andere Funktionen, die heute stärker von Relevanz sind, wie Wissensgenerierung oder Schulentwicklungsfunktion, kaum eine Rolle (nähere Ausführung → Abschnitt 3.3.3).

Im Rahmen dieser Arbeit geht es um die aktuell existierenden Schulinspektionen: Diese sind ab 2004 ungefähr zeitgleich in allen Bundesländern eingeführt worden (vgl. Döbert et al. 2008, S. 137). Dies ist nach Füssel insofern überraschend, als es auf Bundesebene keinen Beschluss gab, „sondern einzig [...] parallellaufende Entwicklungen in einzelnen Bundesländern sich fast virenähnlich über die ganze Republik ausgedehnt“ (Füssel 2008, S. 153) haben. Dabei gibt es in der Literatur verschiedene Begründungen, wie es zu der Einführung von Schulinspektionen gekommen ist: So wird häufig auf die „schlechten“ Ergebnisse Deutschlands in der PISA-Studie verwiesen (z. B. Ehren 2016; Gärtner 2013; Böttcher und Keune 2011; Altrichter und Kemethofer 2016). Jedoch stellen Schulinspektionen nach Döbert et al. (2008) keine unmittelbare Reaktion auf PISA dar, da sie nicht den ersten Post-PISA-Aktivitäten zugerechnet werden (vgl. ebd., S. 137). Dennoch lässt sich festhalten, dass die Frage nach guter Schule unter anderem durch die PISA-Studie vermehrt öffentliche Aufmerksamkeit erhalten hat und parallel dazu das Thema um die Jahrtausendwende in den Erziehungswissenschaften Konjunktur hatte (vgl. Helmke et al. 2018).

Des Weiteren wird die Einführung von Schulinspektionen in der Literatur oft damit begründet, dass Schulen von bildungspolitischer Seite mehr Handlungsspielraum gelassen wurde (vgl. Gärtner 2013, S. 696; Artelt 2007, S. 132). Jene Entwicklung hatte zur Folge, dass sich die Kultusministerien stärker in der Verantwortung sahen, zu kontrollieren, wie Schulen mit der gewachsenen Autonomie umgehen (vgl. Heinrich 2007). Demnach sollte sich bildungspolitische Steuerung stärker darauf konzentrieren, inwiefern von Seiten der Schule Zielvorgaben erfüllt werden (vgl. Landwehr 2011, S. 55) und dies sollte im Fall der deutschen Bundesländer unter anderem von Schulinspektionen überprüft werden. In diesem Zusammenhang kritisiert jedoch Avenarius, dass die Kontrolle von eigenverantwortlichen Schulen eine „*contradictio in adiecto*“ sei (vgl. Avenarius 2006, 6f.). So würde die Etablierung von Kontrollmaßnahmen implizieren, dass Schulen erneut an Autonomie verlieren.

Möglicherweise ist es daher treffender, die beiden oben genannten Begründungen für die Einführung von Schulinspektionen (schlechte PISA-Ergebnisse sowie notwendige Kontrolle von Schulen aufgrund erhöhter Autonomie) nicht als singuläre Auslöser zu verstehen, sondern als Rahmenbedingungen, die die Einführung von Schulinspektionen begünstigt haben. In diesem Zusammenhang sind auch der Trend hin zu evidenzbasierten Entscheidungen im bildungspolitischen Bereich (Heinrich 2015) sowie der Aufstieg der empirischen Bildungsforschung (vgl. Aljets 2015) zu nennen: Zwei Entwicklungen im deutschen Bildungssystem, die sich gegenseitig bedingen. Dabei wird unter evidenzbasierter Bildungspolitik verstanden, wenn „politische Entscheidungen über wissenschaftliches oder zumindest systematisch produziertes Wissen rationalisiert werden können“ (Lambrecht und Rürup 2012, S. 57). Dies basiert auf der Idee, dass mithilfe von empirischen Daten „<evidenzbasierte Bildungspolitik> abseits der festgefahrenen Grabenkämpfe betrieben werden kann“ (Aljets 2015, S. 311; siehe auch Fend 2014, S. 27). Dabei steht die Hinwendung zu evidenzbasierter Bildungspolitik im Zusammenhang mit dem Konstanzer Beschluss, der 1997 von der Kultusministerkonferenz (KMK) verabschiedet wurde (vgl. Kultusministerkonferenz 1997). Damit wurde von Seiten der KMK entschieden an internationalen als auch an nationalen Vergleichsarbeiten teilzunehmen: Mit der Folge, dass

Lernstandserhebungen als eine Form von empirischer Bildungsforschung in Deutschland an Relevanz gewannen.

Im Kontext von evidenzbasierter Bildungspolitik und dem Aufstieg empirischer Bildungsforschung kann auch die Einführung von Schulinspektionen gesehen werden. So orientieren sich die aktuell bestehenden Schulinspektionen – im Gegensatz zu früheren Inspektionsformaten in Deutschland (Brüggemann et al. 2011) – bei der Datenerhebung an empirischen Methoden der Sozialforschung. Die erhobenen Daten sollen unter anderem den Zweck erfüllen, bildungspolitischen Akteuren als Entscheidungsgrundlage zu dienen. In Bezug auf evidenzbasierte Bildungspolitik ist jedoch kritisch zu betrachten, inwiefern im Fall von Schulinspektionen davon gesprochen werden kann, dass die Datenerhebung zur Generierung von „Evidence“² dient. So wird in der Literatur unter Evidence ein Beleg für die Wirksamkeit einer Handlung verstanden (vgl. Bellmann und T. Müller 2011, S. 22), d. h. die Messung davon, ob schulische Prozesse erfolgreich sind. Demnach liegt es nahe im Fall der Schulinspektionen von „systematischer Wissensgenerierung“ (vgl. Dietrich und Lambrecht 2012, S. 60) zu sprechen anstatt von der Generierung von Evidence im klassischen Sinne.

Darüber hinaus stellt das Vorgehen der Schulinspektionen eine Alternative zu Studien, wie PISA, TIMSS oder IGLU dar, die üblicherweise gemeint sind, wenn vom Aufstieg der „empirischen Bildungsforschung“ gesprochen wird: Im Gegensatz zu den Lernstandserhebungen legen die Schulinspektionen den Fokus auf die Prozess- und nicht auf die Outputebene (vgl. Maritzen 2008, S. 93; siehe auch Schwartz et al. 2016, S. 7; Landwehr 2011, S. 56; Lambrecht 2013, S. 226). Demnach können sie aus bildungspolitischer Perspektive als Ergänzung zu jenen Studien gesehen werden, da sie in Bezug auf noch fehlende Daten eine Lücke schließen sollen. Des Weiteren wird die Datenerhebung bei den Schulinspektionen nur zum Teil von wissenschaftlichem Personal realisiert und die Schulinspektionen sind als durchführende Instanz nicht Teil des Wissenschaftssystems, sondern der Bildungsadministration (nähere Ausführung → Abschnitt 3.3). Darüber hinaus kritisiert Heinrich in Bezug auf Lernstandserhebungen, dass durch diese ein bestimmtes Verständnis von Bildungsgerechtigkeit im Sinne von Verteilungsgerechtigkeit dominieren würde und andere Verständnisse von Schulqualität, z. B. Teilhabegerechtigkeit oder Anerkennungstheorie in der öffentlichen Diskussion kaum stattfinden (vgl. Heinrich 2010, S. 47). In diesem Zusammenhang stellt sich die Frage, inwiefern sich die Schulinspektionen in ihrem Verständnis von Schulqualität von den Lernstandserhebungen abgrenzen. Dies kann im Rahmen der empirischen Analyse in dieser Arbeit zum Teil herausgearbeitet werden.

Abschließend ist kritisch zu hinterfragen, inwiefern die vorhandenen empirischen Daten – seien es die Lernstandserhebungen oder die Daten der Schulinspektionen – tatsächlich als Grundlage für bildungspolitische Entscheidungen genutzt werden (vgl. Wiseman et al. 2010). So gibt es in Bezug auf Schulinspektionen wenig verfügbares Wissen zu der Frage, ob und wie jene Daten von bildungspolitischer Seite genutzt werden. Dieser Aspekt wird im Folgenden nicht erneut aufgegriffen, da der Fokus dieser Arbeit auf der Entwicklung der Messinstrumente und

2 In Bezug auf den Begriff der evidenzbasierten Steuerung ist zu berücksichtigen, dass im englischsprachigen Bereich unter „evidence-based“ etwas anderes verstanden wird als im Deutschen unter „Evidenz“: „<Evidence> meint eben nicht das augenscheinlich auf der Hand Liegende, sondern die durch empirische Forschungsmethoden gewonnenen Belege oder Hinweise, die für die Wirksamkeit einer bestimmten Methode oder Maßnahme sprechen.“ (vgl. Bellmann und T. Müller 2011, S. 11). In diesem Zusammenhang argumentiert Bellmann, dass im deutschsprachigen Bereich auch Formen des Bildungsmonitorings, die keinen Anspruch auf kausale Zusammenhänge zwischen einzelnen Indikatoren erheben, der evidenzbasierten Steuerung zugeschrieben werden (vgl. ebd., S. 22). So würden viele Forschungsprojekte für sich in Anspruch nehmen, evidenzbasierte Forschung zu machen, aber „den strengen Kriterien des Paradigmas nur zum Teil oder auch gar nicht genügen“ (ebd., S. 22).

nicht auf der Frage liegt, welche Wirkung die Daten haben, die mithilfe der Messinstrumente erhoben wurden.

3.3 Schulinspektionen im Bildungssystem

In Bezug auf die Entwicklung der Messinstrumente liegt es nahe, im Folgenden die strukturelle Verankerung von Schulinspektionen im Bildungssystem näher zu betrachten. So sind die Handlungslogiken von Schulinspektionen bzw. deren Mitarbeiter_innen möglicherweise davon geprägt, welche Beziehungen zu anderen Akteuren im Bildungssystem bestehen und welche Funktionen die Schulinspektion im Bildungssystem erfüllen soll.

3.3.1 Educational Governance-Ansatz

Bei der Betrachtung des situativen Kontexts von Schulinspektionen bietet es sich an auf den Educational Governance-Ansatz als sensibilisierende Forschungsperspektive (vgl. Maag Merki et al. 2014; siehe auch Dietrich 2018) zurückzugreifen. Dabei stellt dieser eine wertvolle Ergänzung zum Sensemaking-Ansatz dar, welcher im Rahmen dieser Arbeit als theoretischer Rahmen dient, um die Deutungen von Inklusion und die Entwicklung der Messinstrumente abbilden zu können (nähere Ausführung → Kapitel 5). Beiden Ansätzen ist gemein, dass sie sich davon distanzieren, dass Steuerung nach dem Top-Down-Verfahren funktioniert und formal bestehende Hierarchien als Erklärung für Handlungen dienen (in Bezug auf Educational Governance-Ansatz, siehe Kussau und Brüsemeister 2016, S. 21). Des Weiteren können durch den Educational Governance-Ansatz sowohl individuelle als auch kollektive Akteure in den Blick genommen werden (vgl. ebd.). Dabei kennzeichnet sich das erweiterte Steuerungskonzept der Educational Governance durch folgende Sichtweise auf die Realität:

„Die Perspektive von Educational Governance lenkt die Aufmerksamkeit der empirischen Bildungsforschung darauf, dass Leistungen der Bildungssysteme nicht von einem, sondern von vielen Akteuren hergestellt werden, die durch Interdependenzen voneinander abhängig sind und Akteurskonstellationen ausbilden. Im Gegensatz zum Steuerungsansatz und unilateralen Maßnahmen ist die Handlungskoordination im Mehrebenensystem der Fokus der Analyse“ (ebd., S. 44).

Demnach distanziert sich der Educational Governance-Ansatz von einer „linearen Steuerbarkeit“ (Dietrich 2018, S. 75) im Bildungssystem und nimmt die „spezifischen Interessen und Handlungslogiken“ (ebd., S. 76), die auf den einzelnen Ebenen des Bildungssystems existieren, in den Blick. Diese Perspektive ist in Bezug auf die Arbeit insofern von Relevanz, als dass es darum geht das Handeln innerhalb von Schulinspektionen zu verstehen.

Als Ergänzung bietet es sich an, die theoretischen Überlegungen von Dietrich (2018) zu berücksichtigen. Mit dem Begriff der „rekonstruktiven Governanceforschung“ schlägt dieser eine Reformulierung des Educational Governance-Ansatzes vor (siehe auch Lambrecht 2018): Danach ist die bisherige Forschung, die auf diesem Ansatz basiert, häufig geleitet von der „Divergenz zwischen <Soll> und <Sein>“ (Dietrich 2018, S. 80). Es fehlt demnach an einer kritischen Analyse darüber, welche normativen Prämissen bei der Durchführung von Studien von Seiten der Forscher_innen eingenommen oder übernommen werden. So müsste es nach Dietrich stärker darum gehen „eine Rekonstruktion der Bedeutung der Sinnstiftung und Ausgestaltung des jeweiligen Reforminstruments hinsichtlich der interessierenden Ausformung der Akteurskonstellation“ (ebd., S. 80) herzustellen. Angewandt auf das Vorgehen in dieser

Arbeit bedeutet dies, dass die Frage danach, ob die Implementation von Inklusion als politische Vorgabe gelungen bzw. gescheitert ist, nicht im Vordergrund steht, sondern, dass es primär darum geht zu verstehen, wie es zu den jeweiligen Deutungen gekommen ist. Des Weiteren impliziert dies, dass die normativen Setzungen, die durch die Autorin dieser Arbeit vorgenommen werden, kritisch reflektiert werden (nähere Ausführung → Kapitel 6).

3.3.2 Schulinspektion als Teil der öffentlichen Verwaltung

Schulinspektionen sind Teil der Bildungsadministration in den jeweiligen Bundesländern und somit ein Akteur der öffentlichen Verwaltung. Dies ist abzugrenzen von anderen Akteuren im In- und Ausland, die die Qualität von Schulen oder Bildungseinrichtungen messen. So werden Schulratings in den USA beispielsweise von Immobilienfirmen veröffentlicht (vgl. Strauss 2017) oder im Fall von Deutschland werden Schulen unter anderem durch Stiftungen evaluiert, z. B. der alljährliche „Deutsche Schulpreis“ der Robert Bosch Stiftung (2018). Im Vergleich zu Schulinspektionen bestehen bei diesen Akteuren aufgrund der strukturellen Verankerung andere Abhängigkeiten bzw. liegt es nahe, dass diese mit der Messung von Qualität zum Teil auch andere Interessen verfolgen.

Dagegen nehmen Schulinspektionen als Teil der öffentlichen Verwaltung innerhalb des deutschen Bildungssystems die Rolle eines intermediären Akteurs (vgl. Brüsemeister 2007, S. 8) ein: So befinden sich im Mehrebenensystem auf höherer Ebene die Bildungspolitik und unterhalb der Schulinspektion die Schulebene (vgl. ebd., S. 8):

„Sie tragen zur Systemkoordination bei, indem sie zwischen den Akteuren verschiedener Ebenen vermitteln insbesondere zwischen den regulativen Vorgaben der Bildungspolitik und -administration einerseits und den Schulen und Klassenzimmern als den Orten der Erstellung der Systemleistung andererseits“ (Altrichter und Kemethofer 2016).

Neben der Schulinspektion ist die Schulaufsicht ein weiterer Akteur auf der intermediären Ebene. Der Schulaufsicht werden im deutschen Bildungssystem „als weisungsbefugte oder gar disziplinierende Instanz und Rechtsberatung“ (Lohmann und Reißmann 2007, S. 15) verschiedene Aufgaben zugeschrieben, z. B. Kontrolle von Gesetzesvorgaben, Beratung zu Qualitätsfragen sowie Personalangelegenheiten. Demnach liegen Fragen nach Schulqualität auch im Kompetenzbereich der Schulaufsicht. Darüber hinaus argumentiert Füssel nach einer Analyse der schulrechtlichen Regelungen, dass das Handeln der Schulinspektionen ein „Bestandteil der Schulaufsicht“ (vgl. Füssel 2008, S. 156) darstellt: So behält die Schulaufsicht die Aufgabe der Rechts- und der Dienstaufsicht (vgl. ebd.). Durch die Schulinspektion wird die Schulaufsicht „von Teilen ihrer bisherigen Aufgaben – der Schulevaluation – entlastet und [...] dieser Arbeitsaspekt der Schulaufsicht [wird, ap] verselbstständigt, aufgewertet und professionalisiert“ (Rürup 2008, S. 469). Heinrich beschreibt die Aufgabenübernahme von Seiten der Schulinspektion als „funktionale Ausdifferenzierung“ (Heinrich 2015, S. 787). Dahinter steht die Idee, dass die Arbeit der Schulinspektion das Handeln der Schulaufsicht unterstützt: „Polemisch fragen lässt sich mithin, ob es bei den Schulinspektionen um eine Art von <Hilfsbeamten der Schulaufsicht> handelt“ (Füssel 2008, S. 157). Daran anknüpfend gibt Heinrich nach der Analyse von Interviews mit Mitarbeiter_innen der Schulaufsicht zu bedenken:

„Aufgabenbezogen ließe sich kritisch rückfragen, ob es sich streng genommen um eine <funktionale Ausdifferenzierung> oder nicht vielmehr um eine <Verdopplung> von Schulaufsicht handelt, d. h. die datengestützte Schulaufsicht in Form von Schulinspektion als neue Variante gemäß einer funktionalen

Ausdifferenzierung zwischen <datenbasierter> und <nicht-datenbasierter> Schulaufsicht“ (Heinrich 2015, S. 788).

So arbeitet er in der Analyse von Interviews mit Mitarbeiter_innen der Schulaufsicht heraus, dass diese sich kritisch gegenüber dem von Seiten der Schulinspektion generierten Wissen positionieren (vgl. ebd.) und bei der Arbeit mit Schulen kaum auf deren Daten zurückgreifen. In diesem Zusammenhang stellt sich die Frage, wie es – parallel zur Schulaufsicht – zur Einführung von Schulinspektionen gekommen ist: So argumentiert unter anderem Heinrich, dass die Schulaufsicht

„im evidenzbasierten Paradigma in eine Legitimationskrise [geriet, ap], da ihr Handeln sich im Rahmen der damit vorliegenden Datenkonstellation (Monitoring, Schulstatistik, Leistungsdaten etc.) nicht entsprechend der nunmehr maßgeblichen Indikatoren und Outputmaße rechtfertigen ließ“ (ebd., S. 781).

So wurden neue Anforderungen formuliert, die die Schulaufsicht in den gegebenen Strukturen und Praktiken nicht erfüllen konnte: „Vom Augenschein zur Empirie, vom Vorgesuwusstsein zum Nachgewiesenen, vom Bescheidwissen zum Draufschaun, von kasuistischen Vorgehensweisen zu standardisierten Prozessen“ (Maritzen 2008, S. 93). Darüber hinaus argumentiert Maritzen, dass für die Einführung der Schulinspektion sprach, dass die Schulaufsicht „in entscheidenden Qualitätsfragen von <durchschlagender Wirkungslosigkeit> gekennzeichnet“ (ebd., S. 88; siehe auch Stryck 2000, S. 112) war und diagnostiziert somit eine „Krise der Schulaufsicht“ (Maritzen 2008, S. 88; siehe auch Altrichter und Kemethofer 2016). Nach Maritzen kann die Einführung von Schulinspektionen daher von Seiten der Schulaufsicht „durchaus als Machtverlust, mindestens aber als bedrohlich empfunden werden“ (Maritzen 2006, S. 25).

Abschließend lässt sich festhalten, dass die Inspektion in Bezug auf ihren zugewiesenen Aufgabenbereich Überschneidungen zu dem Handeln der Schulaufsicht aufweist. Des Weiteren ist die Schulinspektion formell der Leitungsebene der jeweiligen Kultusministerien unterstellt und in der Hierarchie über der Einzelschule anzusiedeln. In Anlehnung an den Educational Governance-Ansatz können aus diesen formal bestehenden Hierarchien und den sich daraus ergebenden Abhängigkeiten nicht automatisch bestimmte Schlussfolgerungen in Bezug auf das Handeln der Schulinspektionen abgeleitet werden (nähere Ausführung → Kapitel 8). Welche Bezüge zu der Leitungsebene sowie zur Einzelschule bestehen, wird deutlich, wenn im Folgenden näher auf die Funktionen von Schulinspektionen eingegangen wird.

3.3.3 Funktionen der Schulinspektion

In der wissenschaftlichen Literatur gibt es eine Vielzahl von Funktionen, die Schulinspektionen zugeschrieben werden (vgl. Landwehr 2011; Maritzen 2008; für einen Überblick siehe Heinrich et al. 2014). Nach Dederling können die unterschiedlichen Funktionen einem „übergeordneten Ziel zugeordnet werden: der Beförderung einer besseren Schule im Sinne der Steigerung der Qualität“ (Dederling 2012, S. 70). Zur Darstellung der einzelnen Funktionen wird im Folgenden auf das Modell von Landwehr zurückgegriffen, da dieser näher auf die Abhängigkeiten und mögliche Wirkungen der einzelnen Funktionen eingeht. Nach Landwehr werden Schulinspektionen vier Funktionen zugeschrieben: Wissensgenerierung, Entwicklungsfunktion, Kontroll- bzw. Rechenschaftslegungsfunktion sowie Normendurchsetzung (vgl. Landwehr 2011).

Wissensgenerierung: Als Funktion von Schulinspektionen benennt Landwehr die Wissensgenerierung: So erhalten die Schulen datengestütztes Wissen über den Qualitätszustand der eigenen Schule (vgl. Landwehr 2011, S. 39). Der Prozess der Wissensgenerierung dient als Grundlage für weitere Funktionen von Schulinspektionen. Landwehr argumentiert, dass es insofern eine eigenständige Funktion sei, da es den Schulen hilft „den Ist-Zustand besser zu verstehen und bestimmte Problemlagen künftiger fundierter zu thematisieren“ (ebd., S. 39). In diesem Zusammenhang würde es seines Erachtens zu kurz greifen, wenn es nur als relevante Wirkung angesehen werden würde, wenn Schulen für sie bislang unbekanntes Wissen erhielten. In der Praxis würde sich zeigen, dass viele Problemstellungen bzw. auch Stärken den Schulen bereits bekannt seien. Die entscheidende Wirkung jener Funktion sei, dass es zu einer „Offzialisierung einer Problemwahrnehmung“ (ebd., S. 44) kommen würde. Das Wissen über die Qualität der Schule bekommt demnach – unabhängig davon, ob es bereits bekannt ist oder nicht – durch die Schulinspektion einen anderen Status. Diese Funktion ist in Bezug auf die Arbeit von besonderer Relevanz, da es hier darum geht, welche Form von Wissen erfasst wird und wie es zu den Entscheidungen darüber kommt, welches Wissen von Relevanz ist.

Impulsfunktion: Als zweite Funktion von Schulinspektionen benennt Landwehr (2011) die so genannte „Entwicklungsfunktion“. Diese basiert auf der Idee, dass die Daten der Schulinspektion „wirksame Impulse [...] für die Weiterentwicklung der Schule und für die Qualitätsoptimierung der institutionellen Rahmenbedingungen, der Prozesse und der damit erzielten Ergebnisse“ (ebd., S. 48) geben. Damit geht die Hoffnung einher, dass Schulinspektionen durch ihr Handeln nicht nur Informationen über die Schulqualität bieten, sondern auch Handlungsprozesse bei Schulen initiieren (vgl. Dietrich und Lambrecht 2012).

Im Folgenden wird diese Funktion als „Impulsfunktion“ bezeichnet. Dies lässt sich damit begründen, dass der Begriff der „Entwicklung“ irreführend ist: So können auch durch die Kontrollfunktion der Schulinspektion Entwicklungsprozesse bei Schulen ausgelöst werden, aber hierbei wird von einem anderen Steuerungskonzept ausgegangen: So entspricht das Entwicklungsverständnis der Impulsfunktion „dem Paradigma der selbstgesteuerten Organisations- und Praxisentwicklung, wie sie im Bottom-Up-Ansatz der Organisationsentwicklung postuliert wird“ (Landwehr 2011, S. 58). Demnach basiert die Impulsfunktion auf der Idee, dass „Steuerung durch Einsicht“ (vgl. Böttger-Beer und E. Koch 2008; siehe auch Dietrich 2016) funktionieren würde. Wohingegen das Entwicklungsverständnis der Kontrollfunktion (nähere Ausführung folgt) „dem Paradigma der hierarchisch gesteuerten Entwicklung“ (Landwehr 2011, S. 58) entspricht. Danach sollen Schulen dazu angeleitet werden, Schulentwicklungsprozesse zu initiieren und zu realisieren.

In diesem Zusammenhang weist Heinrich auf den Widerspruch hin, der sich daraus ergibt, dass die Schulen in fast allen Bundesländern verpflichtet sind an der Schulinspektion teilzunehmen: Er spricht hier von einem „performativen Selbstwiderspruch, in den sich eine Bildungspolitik begibt, wenn sie versucht eine evidence-based policy per Erlassweg durchzusetzen, d. h. mit dem Oxymoron einer <verordneten Einsicht> operiert“ (Heinrich 2015, S. 786). Dieser Widerspruch lässt sich dadurch erklären, dass den Schulinspektionen im Bildungssystem diverse Funktionen zugeschrieben werden. So ergibt sich aus der Kontrollfunktion, die im Folgenden dargestellt wird, die Notwendigkeit, dass alle Schulen evaluiert werden müssen.

Kontroll- bzw. Rechenschaftsfunktionsfunktion: Als dritte Funktion von Inspektionen beschreibt Landwehr (2011) die Kontroll- bzw. Rechenschaftsfunktionsfunktion. Während die

Impulsfunktion auf der Annahme basiert, dass es immer Anlässe zur Optimierung gibt, entsteht nach Power die Notwendigkeit etwas zu kontrollieren, wenn Misstrauen existiert (vgl. Power 1997, S. 1).

In Bezug auf die Kontroll- bzw. Rechenschaftsfunktionsfunktion steht im deutschen Fall die Interaktion zwischen „oben“, d. h. dem Kultusministerium“ und „unten“, der Schule, im Mittelpunkt. Welche Perspektive auf diese Interaktion eingenommen wird, unterscheidet sich in Abhängigkeit davon, ob von der Kontroll- oder der Rechenschaftsfunktionsfunktion gesprochen wird. Ausgehend von der Kontrollfunktion wird „von oben“ sichergestellt, dass die Einzelschule die an sie formulierten Erwartungen erfüllt. Bei der Rechenschaftsfunktionsfunktion ändert sich insofern die Perspektive, als die Schule gegenüber „oben“ Verantwortung für das eigene Handeln übernimmt, Informationen zur Qualität liefert und das eigene Handeln rechtfertigt. Des Weiteren kann die Rechenschaftsfunktionsfunktion auch so interpretiert werden, dass gegenüber „außen“, z. B. der interessierten Öffentlichkeit, ein Einblick in das eigene Handeln gegeben wird (vgl. Landwehr 2011, S. 40).

In diesem Zusammenhang ist zu berücksichtigen, dass die Berichte der deutschen Schulinspektionen jedoch nur begrenzt für „außen“ zugänglich sind. Hierzu bemerkt Biesta (2010, S. 40), dass das stärkere Streben nach Accountability bzw. Rechenschaftsfunktion von vielen mit der Idee verknüpft wird, dass in einem demokratischen Staat die Bevölkerung als relevanter Akteur involviert sein müsste. Dies ist in der Praxis häufig jedoch nicht der Fall. Wie bereits oben erwähnt, werden die Inspektionsberichte nur in drei Bundesländern der interessierten Öffentlichkeit zugänglich gemacht. Ansonsten geht es primär darum, dass die Schulen gegenüber dem Kultusministerium sowie den Schulträgern Rechenschaft über ihr Handeln ablegen. Dabei bleibt es jedoch der Schulinspektion überlassen zu entscheiden, worüber Rechenschaft abgelegt werden muss.

Im Vergleich zur Impulsfunktion äußert Landwehr die These, dass die Kontrollfunktion „die wichtigere Existenzberechtigung“ (Landwehr 2011, S. 58) für Schulinspektionen darstellt. So könnten Entwicklungsprozesse nicht nur durch Impulse von Seiten der Schulinspektionen, sondern auch durch interne Evaluationen angestoßen werden. Dagegen „braucht eine öffentlich-rechtliche Institution Formen der Rechenschaftsfunktion, die einem differenzierteren Qualitätsverständnis und den Glaubwürdigkeitsansprüchen einer aufgeklärten Gesellschaft standhalten“ (ebd., S. 58). Demnach kann die Impulsfunktion auch durch andere Verfahren erreicht werden, wogegen die Kontrollfunktion auf der Grundlage von empirischen Daten ein Alleinstellungsmerkmal der Schulinspektionen darstellt. In der Praxis zeigt sich jedoch, dass sich Schulinspektionen in den letzten Jahren in ihren Verfahrensschritten stärker an der Impulsfunktion orientiert haben und sich zugleich von der Kontrollfunktion distanzieren haben. Jene Tendenz ergab sich nach Heinrich aus der <Krise der Schulinspektion>:

„Diese speist sich vornehmlich aus der Skepsis, inwieweit Schulinspektionen den Steuerungserwartungen der evidenzbasierten Schulentwicklung gerecht werden, die ihnen im Rahmen des Governance-Programms einer evidenzbasierten Steuerung zugeschrieben werden“ (Heinrich et al. 2014, S. 43).

Normen(durch)setzungsfunktion: Diese Funktion basiert auf der Idee, dass sich die Schulen in Vorbereitung auf die anstehende Inspektion mit den Instrumenten auseinandersetzen und durch die Instrumente „Aufmerksamkeitsverschiebungen“ (vgl. Landwehr 2011, S. 60) stattfinden. In der Beschäftigung mit den Instrumenten erfahren die Schulen, was die Schulinspektionen unter einer guten Schule verstehen und welche Anforderungen von Seiten der Kultusministerien an sie formuliert werden. Im nächsten Schritt können sie reflektieren, in-

wiefern sie diese Anforderungen erfüllen. Die Schulinspektion ist demnach ein Instrument, um Normen zu kommunizieren, diese bei den Schulen zu verankern und durch die Inspektion auch deren Einhaltung einzufordern (vgl. Landwehr 2011, S. 40).

Jene Funktion spielt in der Literatur zu Schulinspektionen kaum eine Rolle (vgl. ebd., S. 60; nähere Ausführung siehe Lambrecht 2018). Demnach gibt es kaum Informationen darüber, ob die Normendurchsetzung nachhaltig wirksam ist. Hierbei stellt sich auch die Frage, ob die Schulinspektionen überhaupt überprüfen, inwiefern es zwischen den Inspektionszyklen zu Veränderungen bei den Einzelschulen gekommen ist (ausführliche Diskussion → Kapitel 8).

In Bezug auf diese Arbeit ist die Funktion der Normendurchsetzung insofern von Relevanz, als es in der Arbeit darum geht, welche normativen Erwartungen an Schule formuliert werden und daran anknüpfend, welches Wissen hierfür im Rahmen der Wissensgenerierungsfunktion erhoben wird. In diesem Zusammenhang lässt sich das Funktionenmodell von Landwehr noch erweitern, da die Schulinspektionen nicht nur bestimmte Normen durchsetzen, sondern auch normsetzend agieren, indem sie Instrumente zur Messung von Schulqualität selbst entwickeln. Demnach setzen Schulinspektionen nicht nur bildungspolitische Vorgaben „von oben“ um, sondern fungieren selbst als Policymaker, indem sie diese konkretisieren und inhaltlich ausfüllen.

Funktionsmix: Als Zwischenfazit lässt sich festhalten, dass Inspektionen diverse Funktionen zugeschrieben werden. Maritzen (2006, S. 6) spricht daher von einem Funktionsmix, der sich dadurch kennzeichnet, dass die einzelnen Funktionen sich zum Teil gegenseitig bedingen und andere nur begrenzt miteinander kompatibel sind.

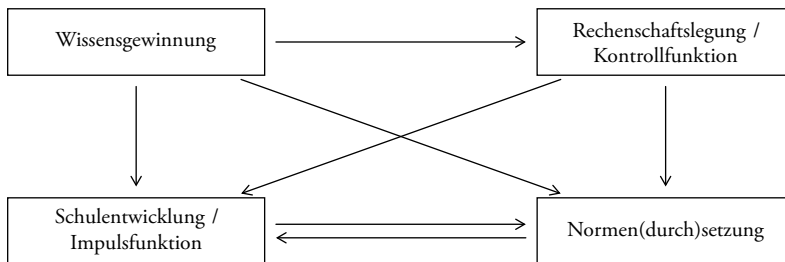


Abb. 3.1: Wirkungsbereiche der externen Schulevaluation (basierend auf Landwehr 2011)

Die bestehenden Abhängigkeiten zwischen den einzelnen Funktionen hat Landwehr (2011) in einem Modell visualisiert (siehe Abbildung 3.1): Beispielsweise dient datengestütztes Wissen zur Schulqualität als Grundlage, um die Kontrollfunktion bzw. die Impulsfunktion zu ermöglichen. In Ergänzung an Landwehrs Modell lässt sich festhalten, dass sich nicht nur die Funktion der Wissensgenerierung auf die Normendurchsetzung auswirkt. Auch die Entscheidung darüber, welches Wissen generiert wird, ist davon abhängig, welche Normen durchgesetzt werden sollen. Darüber hinaus kann sich auch die Impuls- sowie die Kontrollfunktion darauf auswirken, welches Wissen erhoben wird. Wenn beispielsweise im Rahmen der Kontrollfunktion überprüft werden soll, inwiefern bestimmte rechtliche Vorgaben eingehalten werden, spiegelt sich dies auch darin wider, welcher Wissensbedarf bzw. welche Anforderungen an die Schulen formuliert werden.

Darüber hinaus ist in der Darstellung der einzelnen Funktionen sichtbar geworden, dass zwischen den Funktionen auch Spannungsfelder existieren, z. B. ergibt sich aus der Kontroll-

funktion die Verpflichtung die Qualität von allen Schulen zu überprüfen. Dies ist jedoch aus der Perspektive der Impulsfunktion widersprüchlich zu dem Konzept „Steuerung durch Einsicht“. So geht dieses Steuerungskonzept davon, dass sich Inspektion und schulische Vertreter auf Augenhöhe begegnen (vgl. Dietrich und Lambrecht 2012). Dagegen impliziert die Kontrollfunktion ein hierarchisches Verhältnis zwischen den beiden Akteuren. In Bezug auf jene Funktionen nach Landwehr ist zu berücksichtigen, dass diese den Schulinspektionen in der Theorie zugeschrieben werden. Demnach kann es sein, dass es zwischen den Bundesländern bzw. zwischen einzelnen Mitarbeiter_innen von Schulinspektionen variiert, welche Funktionen für relevant erachtet werden und wie diese umgesetzt werden. Des Weiteren kann es auch Unterschiede darin geben, wie die Schulinspektion von Seiten der schulischen Vertreter wahrgenommen werden. Die Gewichtung der einzelnen Funktionen wird in der empirischen Analyse (→ Kapitel 8) thematisiert, da dies auch Einfluss darauf hat, wie Inklusion gemessen wird.

3.4 Zwischenfazit

In Bezug auf die aktuell existierenden Schulinspektionen in Deutschland lässt sich Folgendes festhalten: Sie zeichnen sich dadurch aus, dass sie im Bildungssystem verschiedene Funktionen erfüllen sollen, z. B. Kontroll- und Impulsfunktion. Bei der Generierung von Daten orientieren sie sich an empirischen Methoden der Sozialforschung. Innerhalb der öffentlichen Verwaltung werden sie als intermediärer Akteur beschrieben, der mithilfe von empirischer Datenerhebung – als Ergänzung zur Schulaufsicht – für die Sicherung von Qualität auf Einzelschulebene zuständig ist.

Die Schulinspektionen in den einzelnen Bundesländern unterscheiden sich in Bezug auf einzelne Verfahrensschritte und in den Messinstrumenten, die für die Datenerhebung genutzt werden. Des Weiteren ist noch zu berücksichtigen, dass viele Schulinspektionen in den letzten Jahren bereits wieder abgeschafft wurden (Hessen in 2016, Rheinland-Pfalz in 2015) oder pausieren bzw. ausgesetzt sind (Sachsen seit 2015; Thüringen seit 2015; Bayern seit 2018; Baden-Württemberg seit 2017; Schleswig-Holstein von 2009–2016). Dabei wird die Abschaffung der Schulinspektionen unter anderem damit begründet, dass Kosten eingespart werden müssten (z. B. in Rheinland-Pfalz, siehe Ministerium für Bildung, Wissenschaft, Jugend und Kultur, Rheinland-Pfalz 2015) oder im Fall von Bayern auch mit Lehrkräftemangel (GEW Bayern 2018). Des Weiteren gab es verschiedene Studien (nähere Ausführung → Abschnitt 3.5 in diesem Kapitel), die herausgefunden haben, dass Schulinspektionen nur eine geringe Wirkung auf die Qualitätsentwicklung von Schulen hätten. Auch dies hat den öffentlichen Diskurs geprägt und die Abschaffung von Schulinspektionen beeinflusst. Im Folgenden wird im Rahmen des Forschungsstands unter anderem auf diese Studien zur Wirkung von Schulinspektionen näher eingegangen.

3.5 Forschungsstand zu Schulinspektionen

Parallel zu der Einführung von Schulinspektionen wurden in Deutschland viele Forschungsprojekte initiiert (vgl. SteBis 2017). Daran anknüpfend folgt nun ein Überblick über Studien, die sich mit Schulinspektionen in Deutschland auseinandersetzen sowie internationale Studien, die Anknüpfungspunkte zu der Fragestellung dieser Arbeit bieten.

3.5.1 Schulinspektionen als ein Akteur im Bildungssystem

Im Rahmen des Forschungsschwerpunkts des Bundesministeriums für Bildung und Forschung „Steuerung im Bildungssystem“ (für eine Übersicht siehe SteBis 2017) hat sich eine Vielzahl von Studien mit der Etablierung der Schulinspektionen in Deutschland auseinandergesetzt. Dabei ging es unter anderem um die Frage, wie sich die Schulinspektionen zu anderen Akteuren (z. B. zur Schulaufsicht) verhalten (Heinrich 2015) sowie welche Funktionen den Schulinspektionen im Bildungssystem zugeschrieben werden (z. B. Heinrich et al. 2014; Arbeitsgruppe Schulinspektion 2016; Sowada und Dederig 2016). Da bereits im Abschnitt zur strukturellen Verankerung von Schulinspektionen (→ Abschnitt 3.3) ausführlich auf die Ergebnisse jener Studien zurückgegriffen wurde, wird hier nicht erneut darauf eingegangen.

3.5.2 Wirkung von Schulinspektionen

Des Weiteren gibt es zahlreiche Untersuchungen, die sich in Deutschland sowie international mit der Frage auseinandersetzen, welchen Einfluss Schulinspektionen auf Einzelschulen haben. Es lassen sich hierbei grob folgende Fragestellungen herausgreifen, die im Rahmen von empirischen Studien analysiert werden (für einen Überblick der Forschung siehe Altrichter und Kemethofer 2016; Böttcher und Keune 2010; Husfeldt 2011; Penninckx 2016):

- Inwiefern werden durch Inspektionen Veränderungen bei Schulen initiiert?
- Welche Faktoren können möglicherweise erklären, dass Schulinspektionen (k)einen Einfluss auf Entwicklungsprozesse an der Einzelschule haben?

In Bezug auf die erste Frage argumentieren Gärtner et al.: „Schlussendlich muss sich die Schulinspektion daran messen lassen, ob sie zu positiven Auswirkungen auf das Lernen der Schüler und deren Kompetenzaufbau beiträgt“ (Gärtner et al. 2009, S. 16). Demnach gibt es eine Vielzahl an Studien, die sich damit auseinandergesetzt haben, inwiefern die Durchführung der Schulinspektion dazu geführt hat, dass an Schulen Maßnahmen zur Schulentwicklung getroffen wurden (z. B. Dederig und S. Müller 2011 für Nordrhein-Westfalen; Gärtner et al. 2009 für Brandenburg; F. Huber 2008 für Bayern; Würster und Gärtner 2013; Gärtner et al. 2013 für Berlin und Brandenburg; Böhm-Kasper und Selders 2013 für Mecklenburg-Vorpommern, Thüringen, Baden-Württemberg und Niedersachsen; Diedrich und Pietsch 2013 für Hamburg; Böttcher und Keune 2011 für Hessen; für einen Überblick siehe Altrichter und Kemethofer 2016). Bei einer Untersuchung von Huber (2008) zu der bayerischen Schulinspektion gab über die Hälfte der Schulen an, dass sie nach dem Besuch der Schulinspektion zwischen einer und sechs Maßnahmen realisiert haben. Bei Gärtner et al. (2009) war dagegen die Anzahl derjenigen, die keine Maßnahme nach der Schulinspektion ergriffen haben, im Fall von Brandenburg sehr viel geringer als bei der Studie zu Bayern von Huber (15% statt 43%). Die Studien, die auf Selbsteinschätzungen von Schulen basieren, stehen jedoch in der Kritik. So argumentieren Altrichter und Kemethofer (2016), dass die Schulleitungen bei dieser Art von Befragungen sozial erwünschtes Antwortverhalten zeigen. Dies machen sie unter anderem daran fest, dass die Antworten der Lehrkräfte, wenn diese auch befragt wurden, weniger positiv ausfallen als die Antworten der Schulleitungen. Des Weiteren stellt sich auch die Frage, wie nachhaltig die Maßnahmen sind, die von Seiten der Schulen getroffen wurden, und ob diese zum gewünschten Ziel geführt haben. So ist die reine Anzahl der getroffenen Maßnahmen allein wenig aussagekräftig. In Bezug auf jene Untersuchungen zu Schulinspektionen und Schulentwicklung ist darüber hinaus kritisch zu betrachten, dass diese häufig nicht berücksichtigen, dass eine Vielzahl von

Faktoren beeinflusst, ob Schulentwicklungsprozesse initiiert werden. So sind die Wirkungen der Schulinspektion auf Schulentwicklungsprozesse „nicht einfach linear ableitbar“ (Heinrich et al. 2014, S. 26). Erste Hinweise gibt das Wirkungsmodell von Ehren und Visscher (2006): Danach sind nicht nur diverse Merkmale der Schulinspektion einzubeziehen, z. B. Art des Feedbacks, sondern auch externe Impulse, Reaktionen der Schule auf die Inspektion sowie weitere Merkmale der Schule, z. B. Lernbereitschaft (vgl. ebd., S. 44). Daran anknüpfend haben Behnke und Steins durch die Analyse vorhandener Studien herausgearbeitet, welche Faktoren allein den Feedbackprozess beeinflussen, z. B. Glaubwürdigkeit oder Qualität des Feedbacks (vgl. Behnke und Steins 2017). Demnach verbirgt sich hinter den einzelnen Dimensionen von Ehren und Visscher (2006) eine Vielzahl von unterschiedlichen Ausprägungen. In Bezug auf die Wirkung von Schulinspektionen bedeutet dies, dass andere Faktoren deren behindern, aber auch fördern können. Ergo:

„Die Wirksamkeit der externen Schulevaluationen wird sehr oft gleichgesetzt mit ihrer Entwicklungswirksamkeit (...). Diese Auffassung und die damit signalisierte Ziel- und Funktionsunterstellung ist nicht (...) falsch; ist aber (...) unhaltbar verkürzt“ (Landwehr 2011, S. 37).

Des Weiteren kritisiert Lambrecht, dass bei den Studien zu wenig berücksichtigt wird, welche Empfehlungen die Schulen erhalten, da diese von den tatsächlichen Datensätzen abweichen können (vgl. Lambrecht 2013). Daran anknüpfend gibt es Studien, die sich damit auseinandergesetzt haben, inwiefern Gelingensbedingungen vorhanden sind, damit von Seiten der Schulinspektionen Impulse für Entwicklungsprozesse ausgehen. Dabei werden als Gelingensbedingungen unter anderem das Verfahren der Schulinspektion (vgl. Würster et al. 2016) sowie die Akzeptanz und Rezeption der Schulinspektion durch die Schule (vgl. ebd.; Behnke und Steins 2017; Böttcher und Keune 2010, S. 154) gesehen.

Zusammenfassend hält Heinrich fest:

„Angesichts der (...) internationalen Befunde zur Wirksamkeit des Instruments muss konstatiert werden, dass das Instrument der Schulinspektion politisch nur schwer allein über dessen Wirksamkeit legitimiert werden kann“ (Heinrich et al. 2014, S. 27).

Demnach stellt sich die Frage, ob die Idee noch aufrecht erhalten werden kann, dass Schulinspektionen primär die Aufgabe haben Impulse für Schulentwicklungsprozesse zu geben, weil sie diese Funktion nur begrenzt erfüllen (können). Stattdessen sollten von Seiten der Forschung möglicherweise andere Wirkungsdimensionen wie die Normendurchsetzungsfunktion in den Blick genommen werden (vgl. Lambrecht 2018).

Demnach ist es kritisch zu sehen, dass sich die Forschung zur Wirkung von Schulinspektionen überwiegend auf Schulentwicklungsprozesse fokussiert. So weist Gärtner (2013) darauf hin, dass die Wirkung der Normendurchsetzungsfunktion „bislang kaum systematisch untersucht worden“ (ebd., S. 697) ist. Des Weiteren gibt es auch kein Wissen darüber, inwiefern Evaluationsmaßnahmen wie die Schulinspektion implizieren, dass sich bei den Schulen neue Rechtfertigungsordnungen etablieren: So arbeiten Lowenhaupt et al. (2016) für den US-amerikanischen Kontext heraus, dass Schulleitungen in ihrem Handeln häufig auf Argumentationen, die auf dem Paradigma der evidenzbasierten Steuerung beruhen, zurückgreifen. Dies gibt Hinweise darauf, dass sich evidenzbasiertes Denken und Handeln auch auf der Ebene der Einzelschule durchgesetzt hat. Des Weiteren gibt es im deutschsprachigen Raum auch kaum Forschung dazu, wie sich die Inspektionsergebnisse auf das Handeln der Eltern auswirken, z. B. bei der Schulwahl oder bei der Beteiligung an Schulentwicklungsprozessen.

Für den deutschsprachigen Raum fehlt es auch an Forschung zu der Frage, ob die Inspektionen nicht-intendierte Wirkungen haben (vgl. Bellmann, Duzevic et al. 2016, S. 383). Bellmann et al. können in Bezug auf Formen neuer Steuerung nachweisen, dass sich viele Nebenfolgen evidenzbasierter Steuerung von Bildungseinrichtungen, z. B. Window Dressing, auch in Deutschland zeigen. Hierbei ist jedoch zu berücksichtigen, dass sich die Untersuchung von Bellmann et al. auf diverse Formen bildungspolitischer Steuerung beziehen, z. B. auch Vergleichsarbeiten. In Bezug auf Schulinspektionen gab es in der quantitativen Befragung – nach der Darstellung in Bellmann et al. (2016) – nur zwei Items, die explizit auf Schulinspektionen abzielen (siehe ebd., S. 225):

- „Während der Schulinspektion zeigen Lehrkräfte unserer Schule andere Unterrichtsmethoden als im normalen Schulalltag“ (im Fall von Berlin: 13,6% trifft voll zu; 33,8% trifft eher zu; 42,4% trifft eher nicht zu; 10,3% trifft überhaupt nicht zu).
- „Es kommt vor, dass einzelnen Lehrkräften nahegelegt wird, an den Tagen der Schulinspektion nicht zur Schule zu kommen“ (im Fall von Berlin: 2,0% trifft eher zu; 8,9% trifft eher nicht zu; 89,1 trifft überhaupt nicht zu).

Demnach gibt es kaum Informationen zu der Frage, ob Inspektionen auch nicht-intendierte Wirkungen auf Schulen, z. B. inoffizielle Vorbereitungen durch Schulträger bzw. Schulleitung oder „Anpassungen“ von Dokumenten, haben.

3.5.3 Analyse der Messinstrumente

Während es eine Vielzahl von Studien gibt, die sich mit der Wirkung von Schulinspektionen auseinandersetzen, gibt es nur wenige Untersuchungen zu den Messinstrumenten und den dahinterliegenden Verständnissen von Schulqualität. So vergleichen beispielsweise Ehren und Scheerens (2015), wie die Schulinspektionen in den Niederlanden, England, Österreich (am Beispiel der Steiermark), Tschechien, Irland und Schweden Qualität messen. Dabei liegt der Fokus stärker auf der Operationalisierung und der Frage, ob dies zu validen Ergebnissen führt. Es wird im Rahmen des Artikels jedoch kaum thematisiert, dass es unterschiedliche Verständnisse davon gibt, was unter einer guten Schule verstanden wird.

In Bezug auf die Analyse der Messinstrumente ist des Weiteren die Studie von Osley und Morrison (2000) zu nennen: Sie haben untersucht, inwiefern in den Instrumenten von OFSTED, der Schulinspektion in England, der Abbau von Rassismus thematisiert wird: Demnach gebe es einzelne Indikatoren, die sich auf (Anti-)Rassismus beziehen. Dennoch stellten die beiden Autorinnen fest, dass es an Indikatoren fehlen würde und empfahlen dem Bildungsministerium, dass das Thema bei OFSTED eine größere Rolle spielen müsste (vgl. ebd., S. 10). Des Weiteren zeigte sich in der Studie, dass die befragten Inspektor_innen angaben, dass es ihnen an Expertise fehlen würde, um (Anti-)Rassismus evaluieren zu können (vgl. ebd., S. 5).

Als weitere Studie ist die Untersuchung von Degenhardt (2008) zu Schulinspektionen in Deutschland zu nennen: Hierbei wurde untersucht, ob Inklusion/Integration als Begriffe in den Instrumenten der Schulinspektionen aufgeführt werden und ob die Situation von Schüler_innen mit Sehbeeinträchtigung durch die Messinstrumente in den Blick genommen wird. Hierbei zeigt sich, dass diese Gruppe von Schüler_innen nur in einem Fall (in Bezug auf Förderschulen in Rheinland-Pfalz) in den Instrumenten explizit genannt wird (ebd., S. 24). Auch bei der Analyse, ob und wie häufig der Begriff „Inklusion“ verwendet wird, stellt Degenhardt fest, dass dieser in den Instrumenten von 2007/2008 nicht auftaucht. Demnach gibt Degenhardts

Arbeit erste wichtige Hinweise darauf, welche Relevanz dem Thema vor der Ratifizierung der BRK beigemessen wird. In Bezug auf Inklusion fehlt es jedoch an einer tiefgehenden Analyse davon, was in den Schulinspektionen unter Inklusion verstanden wird, wenn sie den Begriff verwenden bzw. inwiefern möglicherweise Aspekte von Inklusion enthalten sind, ohne dass die Begriffe Integration/Inklusion als solche verwendet werden. Hier setzt die vorliegende Arbeit an. Dabei erlaubt die Analyse der Instrumente, die nach 2009 entstanden sind, auch Aussagen darüber, inwiefern die BRK in den einzelnen Bundesländern implementiert wurde. In Anbetracht der oben dargestellten Untersuchungen lässt sich demnach festhalten, dass es in Bezug auf Schulinspektionen in Deutschland³ kaum Untersuchungen gibt, die sich mit der normativen Komponente von Messinstrumenten, d. h. den dahinterliegenden Wertvorstellungen, auseinandersetzen.

3.5.4 Entwicklung der Messinstrumente

Darüber hinaus hat sich die bestehende Schulinspektionsforschung kaum mit der Frage beschäftigt, wie die Instrumente der Schulinspektionen entstehen. Nach Bos et al. (2007) orientieren sich viele Schulinspektionen bei der Entwicklung der Instrumente am Ausland, z. B. an den Niederlanden. Darüber hinaus werden Forschungsergebnisse zum guten Unterricht berücksichtigt (vgl. ebd., S. 249; siehe auch Heinrich 2010, 62f.). Dabei liegt der Fokus häufig auf der Schuleffektivitätsforschung (vgl. Lambrecht 2018, S. 97 für Hamburg) und somit auf dem dort vorherrschenden Verständnis von guter Schule (nähere Ausführung → Kapitel 4). Weitere Einflussfaktoren können unter anderem sein: „schulrechtliche Regelungen und Rahmenvorgaben (...), Erwartungen seitens der Gesellschaft und der Abnehmer der Schule (...), pädagogische Ziele und Orientierungen“ (Holtappels 2003, S. 41). In diesem Zusammenhang kommentiert jedoch Böttcher: „Auch die Qualitätstableaus, die die Beobachtung und Datensammlung lenken, sind nur rhetorisch durch Verweise auf vermeintlich sichere Befunde der Schulforschung abgesichert“ (Böttcher 2009, S. 9). Bislang basieren jene Verweise in der Literatur, die Hinweise auf die Herkunft der Messinstrumente geben, nicht auf empirischen Untersuchungen, sondern können eher als Vermutungen bzw. anekdotische Evidenzen gelesen werden. Es fehlt demnach an einer systematischen Analyse davon, wie die Messinstrumente der Schulinspektionen entstanden sind und welche Rahmenbedingungen eine Rolle gespielt haben. Für die empirische Analyse der Messinstrumente bietet es sich an theoretische Arbeiten aus der Bewertungssoziologie zu berücksichtigen, welche im nächsten Kapitel thematisiert werden.

³ Im Rahmen der Recherche wurden auch kaum Untersuchungen zu diesem Aspekt für ausländische Schulinspektionen gefunden, jedoch sind hierzu nur begrenzt Aussagen möglich, weil zu erwarten ist, dass diese wahrscheinlich in der jeweiligen Landessprache veröffentlicht wurden.

4 Messinstrumente aus bewertungssoziologischer Perspektive

Im Alltag kommt es ständig zu Bewertungen: Diese Prozesse unterscheiden sich unter anderem darin, welche Bewertungskriterien angelegt werden, ob die angewandten Kriterien explizit kommuniziert wurden und welche Folgen die jeweilige Bewertung impliziert. Im Rahmen der Arbeit liegt der Fokus auf Bewertungen durch Schulinspektionen. Deren Handeln wird im Folgenden als eine Form der Bewertung verstanden: Die Messinstrumente (z. B. der Unterrichtsbeobachtungsbogen) bzw. die einzelnen Indikatoren, die zur Datenerhebung verwendet werden, basieren auf bestimmten Vorstellungen von guter Schule und die Daten werden dazu genutzt, um Aussagen über die Qualität von Einzelschulen zu treffen.

So erheben Schulinspektionen mithilfe von standardisierten Indikatoren¹ sowohl qualitative als auch quantitative Daten. Hierfür greifen sie auf unterschiedliche Bewertungsinstrumente zurück, z. B. enthalten die Unterrichtsbeobachtungsbögen verschiedene Indikatoren, um die Qualität von Unterricht zu beschreiben, z. B. „Der Umgang zwischen Lehrkräften und Schüler/innen ist wertschätzend“. Da die einzelnen Indikatoren das Ergebnis von kollektiven Aushandlungsprozessen darstellen und als zeit- und raumunabhängig gelten (vgl. Timmermans und Epstein 2010, S. 71), werden sie im Folgenden auch als Standards bezeichnet. Dies ist abzugrenzen von Bewertungsprozessen, die sich dadurch kennzeichnen, dass sich Bewertungskriterien situativ herausbilden und nicht schriftlich festgehalten werden.

Um den Gegenstand der Arbeit näher zu beschreiben und die Relevanz von Indikatoren deutlich zu machen, wird auf theoretische Konzepte aus der Soziologie des Bewertens zurückgegriffen (vgl. Lamont 2012). So gab es in den letzten Jahren eine Vielzahl von empirischen Studien zu bewertungssoziologischen Fragestellungen, z. B. zu Bewertungsprozessen im Wissenschaftssystem (z. B. Franzen 2015; Espeland und Sauder 2007; Hamann 2016); in Bezug auf Dienstleistungen im privaten Bereich (Jeacle und Carter 2011), zum Equity – Begriff bei PISA (vgl. Gorur 2014) oder zu Entscheidungsverfahren über UNESCO-Kulturstätten (vgl. Schäfer 2016). Dabei wird in bewertungssoziologischen Arbeiten unter anderem untersucht, welche Implikationen Bewertungen in bestimmten Settings haben, wie sich im Rahmen von kollektiven Aushandlungsprozessen implizit und explizit Bewertungskriterien herauskristallisieren und wie diese begründet werden. Hierfür wird in der Literatur zu Bewertungsprozessen auf verschiedene theoretische Arbeiten zurückgegriffen, z. B. auf Deweys Theorie der Bewertung (vgl. Dewey 1939); auf Durkheim und Mauss' Arbeiten zu Klassifizierung (vgl. Durkheim und Mauss 1987) sowie auf Simmels (1900) Arbeiten zur Philosophie des Geldes. Des Weiteren wird in den letzten Jahren vermehrt auf konventionsoziologische Arbeiten aus der so genannten „französischen“ Schule verwiesen. Diese hat ihren Ursprung am Institut national de la statistique et des études économiques (INSEE) in Paris. Dort haben sich verschiedene Forscher, u. a. Bourdieu, Boltanski und Thévenot, mit den dort entwickelten und verwendeten Indikatoren kritisch auseinandergesetzt und herausgearbeitet, wie diese legitimiert werden (vgl. Diaz-Bone

¹ In der Literatur wird Indikator als Begriff häufig verwendet, um quantitative Größen zu erheben (vgl. Döbert 2009, S. 15). Dennoch wird in der Praxis z. B. beim Human Rights Monitoring sowohl auf quantitative als auch auf qualitative Indikatoren zurückgegriffen.

2015, S. 34). So kann auch der Aspekt der Legitimation sich auf die Auswahl von Indikatoren auswirken und wird daher im Rahmen des Kapitels thematisiert (→ Abschnitt 4.2.3).

4.1 Merkmale von Indikatoren

Um die Bedeutung von Indikatoren beschreiben zu können, wird zunächst der Prozess der Klassifizierung dargestellt und daran anknüpfend wird beschrieben, in welchen Fällen Indikatoren für Bewertungen genutzt werden. Des Weiteren wird näher auf Indikatoren eingegangen, die zur Erhebung von numerischen Werten – also quantitativen Daten – genutzt werden. Diese haben in Bezug auf die Produktion und Kommunikation von Daten besondere Implikationen. Abschließend wird thematisiert, inwiefern Indikatoren als Grundlage für Vergleiche genutzt werden. Bei der Auseinandersetzung mit diesen Konzepten aus der Soziologie des Bewertens werden immer wieder Bezüge zum Gegenstand der Arbeit hergestellt. In diesem Zusammenhang wird auch thematisiert, inwiefern das Verfahren der Schulinspektion von üblichen Messverfahren im Bildungsbereich abweicht, z. B. durch die Kombination von qualitativen und quantitativen Erhebungsverfahren sowie der Verzicht auf Vergleiche oder Rankings.

4.1.1 Indikatoren als Klassifizierungen

In vielen Bereichen des gesellschaftlichen Lebens werden Indikatoren genutzt, um Wissen über bestimmte Sachverhalte zu erlangen: Beispielsweise erheben statistische Ämter Daten wie die Einwohnerzahl von Städten oder sozioökonomische Daten über die Bevölkerung.

In diesem Zusammenhang stellt sich die Frage, inwiefern das Erheben von Daten mithilfe von Indikatoren bereits eine Form der Bewertung darstellt. Hierbei ist zunächst festzuhalten, dass durch Indikatoren eine Form von Klassifizierung stattfindet. Darunter ist zu verstehen, dass „wir die Dinge zu unterschiedlichen Gruppen zusammenfassen, die durch klare Grenzen voneinander geschieden werden“ (Durkheim und Mauss 1987, S. 172). Dabei geht es nicht nur um das Schaffen von Gruppen, sondern diese Gruppen werden „zugleich auch nach überaus speziellen Beziehungen geordnet“ (ebd., S. 176). Klassifizierungen leisten dabei einen Beitrag dazu, soziale Realität zu ordnen und Unsicherheiten abzubauen (vgl. Beckert und Musselin 2013, S. 4). Dies tun sie, indem sie die „Beziehungen zwischen den Wesenheiten begreifbar, intelligibel [...] machen“ (Durkheim und Mauss 1987, S. 249). Ein Beispiel für Klassifizierungen ist die Einteilung von Berufen in bestimmte Gruppen (vgl. Beckert und Musselin 2013, S. 2). So wird im Fall von Deutschland unter anderem von der Bundesagentur für Arbeit auf die KldB (Klassifizierung der Berufe) zurückgegriffen. Bei dieser Klassifizierung von Berufen sind zwei Attribute von Relevanz: die Zugehörigkeit zu einem inhaltlichen Bereich, z. B. Medizin, und das Anforderungsniveau der Tätigkeit. Beispielsweise befinden sich Ärzt_innen auf der höchsten Anforderungsstufe, während Gesundheitspflegehelfer_innen auf der niedrigsten Stufe der vierstufigen Skala im medizinischen Bereich eingeordnet werden. Diese Form der Klassifizierung hat sich im Fall von Deutschland etabliert und wird kaum in Frage gestellt. Diese Form der Institutionalisierung hat jedoch zur Folge, dass häufig der Eindruck entsteht, dass bestimmte Formen der Klassifizierung naturgegeben seien: „The implicit assumption of realism is that the object exists before it is measured and that its nature is the same from one country to another however it is measured“ (Desrosières 2000, S. 179).

Die soziale Konstruktion von Klassifizierungen wird jedoch sichtbar, wenn unterschiedliche Klassifizierungen zum gleichen Gegenstand existieren: So werden beispielsweise im Rahmen

von nationalen Statistiken bestimmte Aspekte nicht in allen Ländern für relevant erachtet (vgl. ebd.). Demnach sind Kategorien stets „historisch und sozial eingebettet“ (Salais 2008, S. 195) und sozial konstruiert (Durkheim und Mauss 1987, S. 176).

Des Weiteren zeigt sich, dass es bei der Entwicklung von neuen Indikatoren häufig zu Konflikten darüber kommt, welche Form der Klassifizierung angewandt werden soll. „Struggles about classification are struggles about the worth of goods and of actors“ (Beckert und Musselin 2013, S. 8). Indem beispielsweise bei der Klassifizierung der Berufe der Fokus auf das Anforderungsniveau gelegt wird, impliziert dies, dass die Wertigkeit eines Berufs sich darüber definiert: „Each standard and each category valorizes some point of view and silences another“ (Bowker und Star 2000, S. 5; siehe auch Krüger und Reinhart 2017). Espeland und Sauder sprechen in diesem Zusammenhang von „selective accountability, meaning that these assessments hold people or organizations accountable on some dimensions while obscuring other aspects of the processes they measure“ (Espeland und Sauder 2016, S. 7). Demnach sind Klassifizierungen nicht als neutral anzusehen, weil dadurch bestimmte Perspektiven auf einen Gegenstand hervorgehoben und andere unsichtbar werden. Dies kann ethisch relevante Implikationen haben (vgl. Bowker und Star 2000, S. 5): Beispielsweise wenn die Hautfarbe von Menschen statistisch nicht erfasst wird, fehlt es auch an Informationen darüber, inwiefern Menschen mit schwarzer Hautfarbe die Teilhabe an bestimmten gesellschaftlichen Bereichen erschwert wird.

Die Entwickler von Indikatoren bewegen sich nach Desrosieres (2000) häufig in einem Spannungsfeld: Auf der einen Seite sind sie sich dessen bewusst, dass die gewählten Klassifizierungen nicht naturgegeben, sondern konstruiert sind und demnach auch andere Formen der Klassifizierungen möglich wären. Auf der anderen Seite ist es notwendig, dass sie nach außen die von ihnen gewählten Klassifizierungen als die einzig möglichen präsentieren, damit die von ihnen erhobenen Daten nicht in Frage gestellt werden (vgl. ebd., S. 174). Es wird daher selten kommuniziert, dass es alternative Formen der Datenerhebung gegeben hätte. Es wird demnach von Seiten der Entwickler_innen explizit angestrebt, dass Indikatoren als objektiv wahrgenommen werden. Auch wenn häufig der Anspruch erhoben wird, dass mithilfe von standardisierten Indikatoren ein objektives Abbild der Wirklichkeit möglich ist, Indikatoren „do not simply capture <reality>. Instead, through assigning objects, practices and people to a fixed set of categories they influence how they are perceived and, furthermore perform(ed)“ (Krüger und Reinhart 2017, S. 277). Im Folgenden geht es nun um Indikatoren, die nicht nur zur Abbildung von Wirklichkeit, sondern auch für Bewertungsprozesse genutzt werden.

4.1.2 Indikatoren als eine Form der Bewertung

„Um von <Qualität> sprechen zu können, ist die von Entscheidungen abhängige normative Komponente der Urteilsbildung nicht nur logisch unentbehrlich, sondern praktisch ausschlaggebend“ (vgl. Heid 2013, S. 410).

In Bezug auf die normative Komponente von Indikatoren bietet es sich an, den situativen Kontext von Erhebungen näher zu betrachten. Dadurch wird sichtbar, ob und wie Indikatoren für Bewertungen genutzt werden. So lässt sich aus einem Indikator, der Informationen zur Geschlechtszugehörigkeit der befragten Personen erfasst, noch nicht herauslesen, inwiefern dieser Indikator für Bewertungen genutzt werden kann. Der Bewertungsmoment kann **erstens** sichtbar werden, wenn die Informationen aus diesem Indikator mit einem anderen Indikator, z. B. mit der Höhe des Einkommens, verknüpft werden. So ist durch eine Verknüpfung möglich zu berechnen, ob die Höhe des Einkommens unabhängig vom Geschlecht ist. Dabei zielt die

Verknüpfung auf folgende Norm ab: Alle Menschen sollten – unabhängig vom Geschlecht – gleich viel verdienen.

Zweitens kann es sein, dass die Relevanz von Indikatoren in Bezug auf Bewertungen erst sichtbar wird, wenn Erwartungen an die Daten formuliert werden. Bei den PISA-Daten wurde beispielsweise ein Durchschnittswert festgelegt: So sind 500 Leistungspunkte „der festgesetzte Mittelwert auf den PISA-Leistungsskalen und der Durchschnittswert, der von Schülerinnen und Schülern in OECD-Ländern erzielt wird. Diese empirische Mitte definiert das Normale“ (Bloem 2016, S. 43). Die Orientierung an den 500 Leistungspunkten ist daher das Formulieren eines Soll-Werts und führt dazu, dass die Messung von Kompetenzen für Bewertungsprozesse genutzt werden kann.

Drittens, bei manchen sozialen Phänomenen wird kein Soll-Wert formuliert, sondern die Einschätzung darüber, ob etwas positiv oder negativ zu bewerten ist, ergibt sich aus dem Vergleich (vgl. Desrosières 2000, S. 177): So findet auch eine Bewertung statt, wenn betrachtet wird, wie sich Daten zwischen mehreren Erhebungszeitpunkten verändern und in welche Richtung sie sich bewegt haben, z. B. die Kindersterblichkeitsrate im Zeitverlauf.

Abschließend lässt sich daher festhalten: Der normative Aspekt von Bewertungen wird häufig erst bei der Auswertung bzw. Darstellung der Daten sichtbar.

4.1.3 Kommensurierung

In Bezug auf Klassifizierungen wurde bereits dargestellt, dass die Entscheidung für eine bestimmte Form von Klassifizierung impliziert, dass alternative Formen wegfallen. Im Folgenden wird nun auf eine bestimmte Form der Klassifizierung näher eingegangen: die Verwendung von Indikatoren zur Erhebung von numerischen Daten. So spielen diese im Fall von Schulinspektionen an verschiedenen Stellen des Verfahrens eine Rolle. Beispielsweise stellen der Unterrichtsbeobachtungsbogen (außer in Schleswig-Holstein) sowie die Onlinebefragungen von Eltern, Lehrkräften und Schüler_innen eine quantitative Form der Datenerhebung dar. Zusätzlich greifen die Schulinspektionen in vielen Bundesländern auf ein Qualitätstableau zurück, um den Schulen Rückmeldungen zu geben. Auch in diesem Fall enthalten die Inspektionsberichte in vielen Bundesländern numerische Werte (nähere Ausführung zu den Instrumenten der Schulinspektionen → Kapitel 6).

Espeland und Stevens (1998) verwenden das Konzept der „Kommensurierung“, um die Implikationen von quantitativen Indikatoren deutlich zu machen: „Soziale Phänomene oder Konzepte, wie beispielsweise Bildung, Umweltschutz, Lebensqualität oder Einstellungen werden in Zahlen ausgedrückt“ (Bloem 2016, S. 41). Numerische Werte ermöglichen, dass Objekte miteinander verglichen werden können, z. B. in Rankings. So stellt Kommensurierung eine Möglichkeit dar eine Fülle von Informationen durch die Umwandlung in numerische Werte zu reduzieren und zu vereinfachen (vgl. Espeland und Stevens 1998, S. 314).

Prozesse der Kommensurierung stehen jedoch aus diversen Gründen in der Kritik: **Erstens argumentieren Espeland und Stevens, dass häufig die Annahme besteht, dass die Übersetzung in numerische Werte ein neutraler Prozess sei** (vgl. ebd., S. 314).

„Indicators are typically presented as taken-for-granted facts. Yet, indicators are not neutral representations of the world, but novel epistemic objects of regulation, domination, experimentation and critique“ (Rottenburg und Merry 2015a, S. 5).

Wie bereits oben dargestellt, stellen „Zahlen stets ein Produkt sozialer Konstruktion“ (Bloem 2016, S. 39; siehe auch Espeland und Sauder 2007, S. 36) dar. Bloem argumentiert daher, dass

der Eindruck einer „vermeintlichen Objektivität“ (Bloem 2016, S. 39) entstehen würde. **Zweitens stellt Biesta in Frage, welche Daten überhaupt verfügbar sind:** „This is the question whether we are indeed measuring what we value, or whether we are just measuring what we can easily measure and thus end up valuing what we (can) measure“ (Biesta 2009, S. 35). Demnach ist der Prozess der Datenerhebung nicht nur davon beeinflusst, welche Wertüberzeugungen sich durchsetzen, sondern auch welche Daten verfügbar sind. Dies kann wiederum dazu führen, dass Datenverfügbarkeit Einfluss darauf hat, was unter Qualität verstanden wird. Daran anknüpfend ist **im dritten Schritt kritisch zu hinterfragen, ob alle soziale Phänomene stets in numerische Werte umgewandelt werden können.** Espeland und Stevens (1998) führen in diesem Zusammenhang verschiedene Beispiele auf, um deutlich zu machen, wann die Umwandlung in numerische Werte an ihre Grenzen stößt: So beschreiben sie den Fall des Yavapai-Volks, welches sich dagegen wehrte, dass ihnen Land entzogen wird. So ging es darum, eine Geldsumme als Entschädigung festzulegen. Das Volk argumentierte dagegen, dass das Land einen Teil ihrer Identität darstelle und es daher nicht möglich sei, einen numerischen Wert festzulegen. **Viertens stehen Formen der Kommensurierung in der Kritik, weil dadurch der jeweilige Kontext nicht berücksichtigt werden kann** (vgl. ebd., 321f.). Espeland (2015) beschreibt diesen Prozess als das Löschen von Narrativen. Indem Kontextwissen nicht erhoben wird, geht auch Wissen darüber verloren, wie die jeweiligen Werte zustande gekommen sind. Hierbei ist in Bezug auf Schulinspektionen jedoch zu berücksichtigen, dass diese zwar numerische Werte verwenden, aber zugleich noch andere Daten erheben. Durch das Erheben von weiteren Daten haben sie demnach die Möglichkeit Kontextwissen zu erwerben und dieses bei der Auswertung einzubeziehen.

Fünftens stehen Formen der Kommensurierung in der Kritik, weil sie häufig nicht intendierte Wirkungen haben. Dies hängt unter anderem damit zusammen, dass numerische Werte nach Espeland und Sauer leicht zu vermitteln sind und nicht in Vergessenheit geraten (Espeland und Sauder 2007, S. 18). In dem Zusammenhang führt Colyvas als mögliche Reaktionen auf die Veröffentlichung von numerischen Werten unter anderem Redefinitionen und Manipulationen auf (vgl. Colyvas 2012, S. 185).

Abschließend lässt sich festhalten, dass das Erheben von numerischen Werten in Bezug auf die Wahrnehmung und Wirkung, z. B. Anschein von höherer Objektivität oder Verlust von Kontextwissen, mit bestimmten Implikationen einhergeht. Daran anknüpfend wird im Rahmen des empirischen Teils auch thematisiert, welche Messverfahren angewandt werden, um die Umsetzung von Inklusion zu erfassen.

4.1.4 Vergleiche von Entitäten

Im Fall von Schulinspektionen gelten größtenteils die gleichen Indikatoren für alle Schulen (nähere Ausführung zu abweichenden Indikatoren → Kapitel 6).² Demnach könnten Vergleiche zwischen Schulen realisiert werden.

„Vergleiche sind Beobachtungsinstrumente, die zwischen Einheiten oder Ereignissen eine Beziehung herstellen. Sie beruhen einerseits auf der Annahme, dass die verglichenen Einheiten in mindestens einer grundlegenden Hinsicht gleich sind, und setzen andererseits ein Vergleichskriterium voraus, das die Verschiedenheit des (partiell) Gleichen beobachtbar macht“ (vgl. Heintz 2010, S. 164).

Heintz betont darüber hinaus die soziale Relevanz von Vergleichen (vgl. ebd., S. 163). Hierbei ist zwischen zwei Arten von Vergleichen zu unterscheiden: Es gibt Vergleiche, die lediglich

² Im Fall von Berlin gibt es seit November 2017 die Möglichkeit, dass Schulen selbst Indikatoren entwickeln.

Unterschiede abbilden, es fehlt jedoch an Hinweisen, welche Werte besser oder schlechter zu bewerten sind (vgl. Heintz 2010, S. 165), z. B. Verteilung von Augenfarben innerhalb einer Klasse. Vergleiche gewinnen jedoch an Macht, wenn diese eine klare Hierarchisierung vornehmen, z. B. Rankings im Sport. So haben Espeland und Sauder in Bezug auf Hochschulen nachgewiesen, welche Implikationen die Kommunikation von Rankings im Fall von US-amerikanischen Law Schools hat: So führt die Veröffentlichung der Rankings dazu, dass Hochschulen sich unter Druck gesetzt fühlen und diverse Gaming-Strategien anwenden, z. B. Änderung von Zulassungsverfahren (vgl. Espeland und Sauder 2007).

In Bezug auf die Indikatoren von Schulinspektionen lässt sich festhalten, dass die Verwendung von standardisierten Indikatoren impliziert, dass unabhängig von den jeweiligen Rahmenbedingungen die gleichen Anforderungen an alle Schulen formuliert und diese für vergleichbar angesehen werden. Bei der Erhebung von numerischen Werten wäre es theoretisch möglich, dass die Schulinspektionen die Ergebnisse als Vergleiche von Schulen bzw. Rankings veröffentlichen, wenn es sich um Indikatoren handelt, die für alle Schulen obligatorisch sind. In der Praxis zeigt sich jedoch, dass das Handeln der Schulinspektionen primär nicht darauf abzielt, die Schulen miteinander zu vergleichen. So werden die Ergebnisse der Schulen nicht in Form von Rankings veröffentlicht (nähere Ausführung → Kapitel 8). In manchen Bundesländern wie Berlin, Brandenburg und Hamburg sind jedoch insofern Vergleiche zwischen Schulen möglich, als dass sich Eltern und andere Akteure über die Website der Ministerien auf die Inspektionsberichte zugreifen können. Dadurch haben sie die Möglichkeit die Schulen in ihrem Bezirk miteinander zu vergleichen.

Von Seiten der Schulinspektion ist es jedoch nicht der Vergleich zwischen Schulen, der im Fokus steht, sondern die Indikatoren repräsentieren ein bestimmtes Verständnis von guter Schule. Bei der konkreten Durchführung einer Inspektion geht es daher darum, wie die jeweilige Schule im Vergleich zu diesem Idealbild von Schule abschneidet. Auch wenn ein Vergleich zwischen Schulen durch die Schulinspektionen nicht angestrebt wird, zeigt sich in den Arbeiten von Sowada und Dederling jedoch, dass der Vergleich mit anderen Schulen bei den konkreten Durchführungen der Inspektionen durchaus eine Rolle spielt, da die Inspektor_innen ihr Wissen über andere Schulen bei der Bewertung einbeziehen (vgl. Sowada und Dederling 2014). Des Weiteren wäre es möglich, dass der Vergleichsmoment auch eine Rolle spielt, wenn betrachtet wird, wie eine Schule bei verschiedenen Inspektionszyklen abgeschnitten hat. Hierbei zeigt sich jedoch in der Praxis, dass dieser Aspekt eher eine geringere Rolle spielt (nähere Ausführung → Kapitel 8).

Zusammenfassend: Indikatoren stellen eine Form der Klassifizierung dar. Sie sind nicht naturgegeben. Demnach kann es temporal und lokal variieren, welche Indikatoren zur Erhebung von Wissen genutzt werden. Indikatoren sind demnach nicht neutral, weil sie bestimmte Perspektiven auf einen Gegenstand hervorheben und andere unsichtbar machen. Indikatoren können im Folgenden als eine Form der Bewertung angesehen werden, wenn sie dazu genutzt werden Aussagen über die Qualität eines Gegenstandes zu treffen: „From such a view, classification appears not only as fundamental social but also fundamental to the social“ (Krüger und Reinhart 2017, S. 165).

Als weiterer Aspekt ist die Wirkung von Indikatoren, die zur Erhebung von numerischen Werten genutzt werden, zu nennen: So können jene Indikatoren unter anderem zur Folge haben, dass Kontextwissen verloren geht und neue Narrative entstehen. Des Weiteren werden numerische Werte häufig dazu genutzt eine Hierarchisierung vorzunehmen. Sie vereinfachen somit Vergleiche zwischen Entitäten.

4.2 Entwicklung von Bewertungsinstrumenten

Im vorherigen Abschnitt wurde deutlich, wodurch sich Indikatoren kennzeichnen und welche Implikationen mit der Verwendung von Indikatoren in Bezug auf die vermeintlich objektive Darstellung von Wirklichkeit einhergehen. Im Folgenden liegt der Fokus auf der Entwicklung von Indikatoren zur Messung von Qualität.

4.2.1 Forschungsstand

Es gibt wenige Studien, die sich mit der Frage auseinandersetzen, wie Indikatoren zur Qualitätsmessung entstehen (vgl. Mehrpouya und Samiolo 2016, S. 13; siehe auch Chenhall et al. 2013): So hat sich beispielsweise Desrosières mit der Frage auseinandergesetzt, wie sich statistische Denkweisen in Nationalstaaten durchgesetzt haben (vgl. Desrosières 2005) und wie es zur Harmonisierung von Indikatoren auf europäischer Ebene gekommen ist (vgl. Desrosières 2000; siehe auch Salais 2006). Hierbei liegt der Fokus jedoch auf einer Beschreibung davon, wie sich die Indikatoren im Laufe der Zeit verändert haben.

Michèle Lamont (2010) setzt sich mit relevanten Einflussfaktoren bei evaluativen Praktiken auseinander. So analysiert sie mit Hilfe von Interviews und teilnehmenden Beobachtungen das Handeln von Auswahlkommissionen, die Stipendien für Wissenschaftler_innen vergeben. Jedoch geht es im Rahmen ihrer Arbeit nicht um die Entwicklung von standardisierten Indikatoren, sondern um spontane Wertzuschreibungen im Rahmen von kollektiven Aushandlungsprozessen (siehe auch Hirschauer 2015). In diesem Zusammenhang arbeitet Lamont verschiedene Einflussfaktoren heraus, die den Bewertungsprozess beeinflussen, z. B. die fördernde Organisation sowie die Disziplinzugehörigkeit (vgl. Lamont 2010, S. 32). So variiert es zwischen den Disziplinen in welchem Maße Kriterien für Qualität bereits etabliert sind (vgl. ebd., 64f.) und was in den einzelnen Disziplinen unter Qualität verstanden wird (vgl. ebd., 64f.). Darüber hinaus arbeitet Lamont heraus, dass auch die Kenntnisse und Überzeugungen der Kommissionsmitglieder den Prozess beeinflussen (vgl. ebd., 64f.).

Während sich Lamont mit situationsspezifischen Bewertungsprozessen auseinandersetzt, analysiert die Studie von Chenhall et al. (2013), wie standardisierte Indikatoren entstehen. Diese dienen dazu, Entsendeorganisationen von Freiwilligen zu evaluieren. Basierend auf den theoretischen Arbeiten von David Stark (2009) beschäftigen sie sich mit der Frage, welche Debatten die Entwicklung von Indikatoren geprägt haben (nähere Ausführung → Abschnitt 4.2.3): So fand zum einen eine Diskussion darüber statt, inwiefern die spezifischen Rahmenbedingungen der zu evaluierenden Organisationen berücksichtigt werden können oder ob der Fokus stärker auf standardisierten Indikatoren liegen sollte, die für alle Organisationen gelten. Zum anderen spielte auch eine Rolle, inwiefern das Evaluationsverfahren sich auf das tatsächliche Handeln der Organisationen auswirkt bzw. auswirken sollte (vgl. Chenhall et al. 2013). Chenhall et al. gehen kaum darauf ein, wie es zu den oben beschriebenen Debatten gekommen ist, z. B. welche Akteursinteressen bei der Entwicklung der Indikatoren berücksichtigt wurden. So fehlt es bei der Studie an einer systematischen Analyse davon, welche Rahmenbedingungen die Entwicklung der Indikatoren geprägt haben. Dagegen wird bei der Studie von Mehrpouya und Samiolo (2016) stärker in den Blick genommen, in welchem Spannungsfeld sich die Entwickler_innen von Indikatoren bewegen. Hierbei analysieren sie die Entwicklung eines medizinischen Index⁴. Jedoch fehlt es auch in dieser Studie an Informationen darüber, inwiefern beispielsweise Aspekte auf der Individualebene, z. B. persönliche Attribute der Entwickler_innen, zu diesen Spannungsfeldern geführt haben.

Aufgrund der fehlenden theoretischen Arbeiten, die sich mit der Entwicklung von Indikatoren auseinandersetzen, wird im Rahmen der vorliegenden Arbeit auf einen Ansatz aus der Organisationssoziologie zurückgegriffen: dem Sensemaking-Ansatz (nähere Ausführung → Kapitel 5). Hierbei ist jedoch zu berücksichtigen, dass der Sensemaking-Ansatz lediglich Vorannahmen über die Deutung von bislang unbekanntem bzw. mehrdeutigen Situationen, Konzepten oder Objekten – in diesem Falle der Umsetzung von Inklusion – trifft. Demnach trifft der Ansatz keine Aussagen über die spezifischen Bedingungen von Bewertungssituationen. In Ergänzung zum Sensemaking-Ansatz bietet es sich daher an zwei Aspekte von Bewertungsprozessen zu berücksichtigen, die gegebenenfalls Hinweise über relevante Rahmenbedingungen bei der Entwicklung von Indikatoren geben: die Konstellationen von Bewertungen als Analysemodell (→ 4.2.2) sowie die (notwendige) Institutionalisierung von Indikatoren (→ 4.2.3).

4.2.2 Bewertungskonstellationen

Bewertungskonstellationen (vgl. Heid 2013; siehe auch Meier et al. 2016) beschreiben näher, welche Komponenten das Ergebnis einer Bewertung beeinflussen können. In Bezug auf die Fragestellung dieser Arbeit kann die Berücksichtigung von Bewertungskonstellationen darüber hinaus Hinweise darauf geben, welche Rahmenbedingungen bei der Entwicklung von standardisierten Indikatoren prägend wirken.

Bei den Bewertungskonstellationen unterscheidet Heid (2013) zwischen den folgenden Komponenten:

1. wer das Qualitätsurteil trifft (Subjekt der Qualitätszuschreibung);
2. wozu ein Urteil getroffen wird (Zweck der Qualitätsbeurteilung);
3. was beurteilt wird (Gegenstand der Qualitätsbeurteilung);
4. welche Kriterien bei der Beurteilung herangezogen werden (Kriterium);
5. wen das Qualitätsurteil erreicht (Adressat der Qualitätsurteil);
6. in welchem Kontext das Urteil getroffen wird (Randbedingungen eines Qualitätsurteils) (vgl. ebd., S. 411).

In Abhängigkeit von *dem Subjekt der Qualitätszuschreibung* kann die Bewertung variieren: So können beispielsweise Überzeugungen und Wissen der Akteure sowie deren Zugehörigkeit zu einer bestimmten Organisation Einfluss auf die Qualitätszuschreibung nehmen. Des Weiteren kann das Subjekt auch darauf einwirken, welchen *Zweck ein Qualitätsurteil* erfüllen soll. Dieser ist jedoch nicht unabhängig von dem so genannten *Gegenstand*, dem Objekt der Qualitätsbeurteilung. Der Beurteilungsgegenstand ist nicht mit der Gegenstandsbeurteilung zu verwechseln (ebd., S. 412). So kann der gleiche Gegenstand, z. B. eine Einzelschule, unterschiedliche Qualitätsurteile erhalten. Darüber hinaus ist bei einem Qualitätsurteil auch zu berücksichtigen, was unter Schule oder Unterricht verstanden wird (vgl. ebd., S. 413). Auch hier finden Formen der Selektion statt, z. B. wenn bestimmte Aspekte von Schule oder Unterricht, wie die Architektur des Gebäudes, nicht Teil des Qualitätsurteils sind. In diesem Zusammenhang ist als weitere Komponente die *Kriterienabhängigkeit eines Qualitätsurteils* zu nennen (vgl. ebd., S. 417). Bei der Bewertung eines Gegenstandes ist es maßgeblich, anhand welcher Kriterien bzw. Indikatoren ein Gegenstand beurteilt wird. In Bezug auf Kriterien weist Heid darauf hin, dass die Personen, die evaluieren, Entscheidungen bezüglich „Diagnose- und Evaluationskonzeptionen, Messmethoden und statistische Verfahren“ (ebd., S. 410) treffen müssen. Demnach kann das Urteil über eine Schule sehr stark variieren je nachdem, welche Kriterien an eine Schule angelegt werden (Heid 2000, S. 41). Obwohl die Auswahl der Kriterien – im Rahmen dieser Arbeit die

Messinstrumente – bei der Qualitätszuschreibung grundlegend ist, wird dies in der Literatur selten diskutiert bzw. problematisiert (vgl. Heid 2013, S. 413).

Des Weiteren sind die *Adressaten von Qualitätsurteilen*, die das Qualitätsurteil erhalten und „aktiv, selektiv, interpretativ und wertend [...] rezipieren“ (ebd., S. 415) zu nennen. Heid führt zwei Aspekte auf, die die Relevanz von Adressaten deutlich machen: Zum einen wird die Bedeutsamkeit eines Qualitätsurteils erst durch die Mitwirkung der Adressaten sichtbar, z. B. indem diese auf das Qualitätsurteil reagieren. Des Weiteren ist es auch abhängig von den jeweiligen Adressaten, wie bestimmte Qualitätsurteile rezipiert werden. So kann es hier große Unterschiede zwischen einzelnen Adressaten geben.

Als weitere Komponente sind die *Randbedingungen eines Qualitätsurteils* zu nennen. Hierbei lässt sich nach Heid (ebd., S. 420) schwer abgrenzen, was zum situativen Kontext eines Qualitätsurteils gehört und was nicht. Des Weiteren weist er daraufhin, dass die Randbedingungen stark durch die vorher aufgeführten Komponenten eines Qualitätsurteils beeinflusst werden, z. B. führt er als eine Randbedingung die Urteilskompetenz des Subjekts an (vgl. ebd., S. 420).

Das Modell nach Heid lässt sich folgendermaßen auf Schulinspektionen anwenden: Die Personen, die die konkreten Inspektionen durchführen, sind das Subjekt der Qualitätszuschreibung. Des Weiteren ist die Perspektive der Instrumenteentwickler_innen einzubeziehen, da diese die Kriterien entwickeln, die zur Bewertung herangezogen werden. Häufig finden beide Positionen in Personalunion statt. Im vorherigen Kapitel (→ Kapitel 3) wurde ausführlich dargestellt, welche Funktionen (- Zweck der Qualitätsbeurteilung) die Bewertungen der Inspektionen im Bildungssystem erfüllen, z. B. Impuls- oder Kontrollfunktion. Der Gegenstand der Qualitätsbeurteilung ist die Einzelschule und über das Ergebnis der Bewertung werden sowohl die schulischen Akteure als auch die Bildungsadministration (und z. T. die Öffentlichkeit) informiert (-Adressat des Qualitätsurteils). In Bezug auf den Bewertungsprozess können verschiedene Randbedingungen eine Rolle spielen, z. B. Qualifikation der Inspektor_innen oder Akzeptanz der Schulinspektionen von Seiten der Schule. Im Rahmen der empirischen Analyse (→ Kapitel 8) wird dargestellt, inwiefern diese Komponenten bei der Entwicklung von Instrumenten tatsächlich eine Rolle gespielt haben.

4.2.3 Institutionalisierung von Indikatoren

Die nähere Betrachtung von Bewertungskonstellationen hat deutlich gemacht, welche Aspekte bei der Entwicklung von Indikatoren prägend wirken können. Darüber hinaus kann der Grad der Institutionalisierung eine Rolle spielen: Dieser lässt Aussagen darüber zu, in welchem Ausmaß Indikatoren in der Praxis genutzt werden, z. B. als Grundlage für politische Entscheidungen oder Veränderungen bei der evaluierten Organisation, und wie etabliert sie sind. Dabei findet die Institutionalisierung von Indikatoren – wenn sie überhaupt stattfindet – zeitlich nach der Entwicklung von Indikatoren statt. Es ist dennoch möglich, dass dieser Aspekt bereits bei der Entwicklung von Indikatoren antizipiert wird, da Instrumenteentwickler_innen ein Interesse daran haben, dass die von ihnen formulierten Indikatoren in der Praxis Anwendung finden und zeitlich stabil sind.

Forminvestitionen als Bemühung um Institutionalisierung: Wenn bestehende Klassifizierungen sich durchgesetzt haben, d. h. genutzt werden, werden sie – wie oben bereits dargestellt – häufig nicht mehr in Frage gestellt (vgl. Beckert und Musselin 2013, S. 7). Um in Bezug auf Indikatoren von einem hohen Grad von Institutionalisierung sprechen zu können, müssen sie

jedoch eine bestimmte Autorität erlangen. Dies hängt nach Rottenburg und Merry von den privaten und staatlichen Organisationen ab, die involviert sind (vgl. Rottenburg und Merry 2015b, S. 11). Darüber hinaus ist die Institutionalisierung von Indikatoren wahrscheinlich, wenn zuvor Unsicherheit darüber bestand, wie etwas zu klassifizieren ist (vgl. Thévenot 2009). Des Weiteren kann es sein, dass Indikatoren für legitim erachtet werden, wenn unterschiedliche Akteure beim Entstehungsprozess beteiligt waren (vgl. Castel 2009). In Bezug auf die Etablierung von Klassifizierungen bzw. Indikatoren spricht Thévenot (2009) von „Forminvestitionen“: Dabei geht um es „Investitionen in die Gültigkeit (Validität im Sinne einer Anerkennung) einer Form hinsichtlich einer bestimmten (sachlich-sozialen) Reichweite und für eine bestimmte zeitliche Dauer“ (Diaz-Bone 2015, S. 89; siehe auch Thévenot 2009).

In Bezug auf die Schulinspektionen stellt sich die Frage, inwiefern die Indikatoren der Schulinspektionen zur Messung von Schulqualität etabliert sind.³ Hierzu gibt es bislang kaum Studien, jedoch gibt es verschiedene Aspekte, die darauf schließen lassen, dass der Grad an Institutionalisierung eher gering einzustufen ist (nähere Ausführung siehe Piezunka 2019): So spricht für einen geringen Grad an Institutionalisierung, dass die Bundesländer sich darin unterscheiden, welche Indikatoren sie anwenden, um die Qualität von Schule zu messen. Demnach gelten die jeweiligen Indikatoren nur im eigenen Bundesland und es gibt nicht ein einheitliches Set an Indikatoren, das deutschlandweit bzw. international angewandt wird. Ein niedriger Grad an Institutionalisierung lässt sich auch daraus ableiten, dass die Existenzberechtigung von Schulinspektionen im öffentlichen Diskurs (z. B. BLLV 2017) immer wieder in Frage gestellt wird – mit der Folge, dass die Schulinspektionen in manchen Bundesländern in den letzten Jahren abgeschafft wurden (nähere Ausführung → Kapitel 3). Auf der anderen Seite lässt sich argumentieren, dass die Indikatoren der Schulinspektionen zu einem gewissen Grad bereits institutionalisiert sind, da sie in den Fällen, die im Rahmen der Arbeit betrachtet werden, schriftlich fixiert sind, von den jeweiligen Ministerien akzeptiert wurden und in der Praxis Anwendung finden. Des Weiteren kann insofern von einem gewissen Grad an Institutionalisierung gesprochen werden, als sich die meisten der betrachteten Schulinspektionen aktuell bereits im zweiten oder dritten Inspektionszyklus befinden und somit argumentiert werden könnte, dass die Indikatoren bereits etabliert sind.

Der Grad an Institutionalisierung kann in Bezug auf die Entwicklung von Instrumenten zu Inklusion insofern von Relevanz sein, als dass diese nicht zu Beginn, sondern ca. 5-8 Jahre nach der Entwicklung des ersten Sets an Instrumenten eingeführt wurden. Bereits zu diesem Zeitpunkt wurde das Verfahren der Schulinspektionen im öffentlichen Diskurs in Frage gestellt bzw. die ersten Inspektionen wurden abgeschafft. Demnach kann es sein, dass die bestehenden Schulinspektionen bei der Überarbeitung der Messinstrumente bzw. der Entwicklung von Instrumenten zu Inklusion insbesondere das Interesse verfolgt haben, dass das Verfahren bzw. die Messinstrumente zukünftig für das Handeln von einzelnen Schulen und für die Bildungsadministration von höherer Relevanz sind. Dieser Aspekt wird im Rahmen der empirischen Analyse thematisiert.

Im Folgenden geht es nun um die Frage, wie die Auswahl an Indikatoren von Seiten der Instrumenteentwickler_innen begründet wird. Die nähere Betrachtung von Rechtfertigungsordnungen lässt Hinweise darauf zu, welche Rahmenbedingungen die Entwicklung von Instrumenten geprägt hat und gibt Hinweise darauf, welche Wertvorstellungen eine Rolle gespielt haben bzw. welche nicht von Relevanz waren.

³ Im Rahmen eines Beitrags, der in 2019 erscheint, setzt sich die Autorin dieser Arbeit intensiver mit der Frage der Institutionalisierung auseinander und greift hierfür auf andere theoretische Zugänge zurück.

Rechtfertigungsordnungen: Im Folgenden wird das Konzept der Rechtfertigungsordnungen von Boltanski und Thévenot (2007) näher dargestellt. Die beiden Autoren haben untersucht, auf welche Rechtfertigungsordnungen⁴ zurückgegriffen wurde, um Indikatoren, die für die statistische Berichterstattung in Frankreich genutzt werden, zu begründen (vgl. ebd.).

Die Existenz von unterschiedlichen Rechtfertigungsordnungen wird sichtbar, wenn die Entwickler von Indikatoren sich darüber einigen müssen, anhand welcher Werteordnungen die Wertigkeit eines Objekts festgelegt wird (Wagner 2011, S. 272; siehe auch Anagnostopoulos et al. 2016). Boltanski und Thévenot (1999) sprechen in diesem Zusammenhang von „critical moments“, weil in solchen Momenten nicht klar ist, welche Werteordnungen angewandt werden sollen. Dabei kennzeichnen sich diese Momente dadurch, dass die Akteure reflektieren, wie sie die Wertigkeit eines Objekts festlegen. Um sich zu rechtfertigen, benennen sie explizit Werteordnungen, die ihrer Meinung nach angewandt werden sollten. Demnach stellen die Werteordnungen „<Denkmodelle> für (...) reales, alltägliches Urteilen und Handeln“ (Diaz-Bone 2015, S. 139) dar.⁵

Boltanski und Thévenot unterscheiden hierbei zwischen sechs Werteordnungen, die später noch ergänzt wurden⁶ und im Folgendem u. a. am Beispiel „Wertigkeit einer Person“ illustriert werden:

Welt des Markts: „Als legitime Praxis gelten der freie Wettbewerb (die Konkurrenz) unter Absehung der jeweiligen Person und der individuell geldvermittelte Tausch (Geld gegen Güter) zwischen Individuen“ (ebd., S. 141). Demnach bemisst sich die Wertigkeit von Personen beispielsweise anhand „ihrer Kaufkraft (Käufer) sowie der Behauptung im Wettbewerb (Verkäufer)“ (ebd., S. 141). Als Werte benennt Diaz-Bone Egoismus und Opportunismus (vgl. ebd., S. 141).

Welt der Industrie: „Personen stehen hier nicht in Tauschbeziehungen zueinander, sondern in <funktionalen> Beziehungen. An die Stelle des Eigeninteresses, das in der Marktwelt legitim war, treten die planerische Rationalität und die (wissenschaftliche) Optimierung der Planung“ (ebd., S. 142). Nach Diaz-Bone hat jene Form der Weltordnung eine „enorme Verbreitung, denn die Strategien der Planung, Standardisierung, Effizienzsteigerung und Verzifferung zur Koordination von Produktion finden sich in vielen sozialen Bereichen“ (ebd., S. 142). Die Wertigkeit eines Objekts wird demnach anhand deren Funktionalität begründet.

Welt des Hauses: „In einer häuslichen Welt richtet sich die Wertigkeit eines Menschen nach seinem Rang in einer Vertrauenshierarchie, der wiederum eine Kette persönlicher Abhängigkeiten zugrunde liegt“ (Boltanski und Thévenot 2011:58 zit. in ebd., S. 143). „Personen sind in dieser Welt durch Verwandtschaftsbeziehungen, personale Abhängigkeit, Hierarchie, Loyalität, gemeinsame Herkunft, Abstammungen und Vertrauen miteinander verbunden.“ Hierbei bemisst sich der Wert nicht an der Orientierung an Standards (Welt der Industrie), sondern

4 Diaz-Bone weist darauf hin, dass eine Vielzahl an Begrifflichkeiten für das gleiche Konzept verwendet wird, z. B. „Werteordnung“, „Rechtfertigungsordnung, Welt (cités), Qualitätskonvention oder Polis“ (Diaz-Bone 2015, S. 140). So betont beispielsweise der Begriff „Welt“, dass in der Koordination für die Akteure in Situationen ein praktisch geteilter Sinnhorizont zur Verfügung steht, der intersubjektiv geteilte (und objektgestützte) Interpretationen und Kognitionen ermöglicht“ (ebd., S. 140).

5 In Bezug auf die Soziologie der Konventionen betont Diaz-Bone, dass jenes Konzept von Boltanski und Thévenot „eine der genuinen Leistungen“ (ebd., S. 135) darstellt, die die Soziologie der Konventionen maßgeblich geprägt hat. Jedoch haben Boltanski und Thévenot damals noch nicht das Konzept der Konventionen verwendet (ebd., 135f.).

6 Als weitere Werteordnung führen unter anderem Thévenot und Boltanski noch die ökologische Konvention sowie die projektbasierte Konvention ein (für ausführliche Beschreibungen siehe ebd.).

die Einzigartigkeit von Menschen und ihre „individuelle Könnerschaft“ (Diaz-Bone 2015, S. 143) werden anerkannt. „Das Gemeinwohl besteht hier in der Stärke der Gemeinschaft (...), die sich ihrerseits um ihre Mitglieder <sorgt> und diese schützt“ (ebd., S. 144).

Welt der Meinung: „Die Wertigkeit ist das Ausmaß an Ruhm, Ehre, Anerkennung und Bekanntheit“ (ebd., S. 144). In der Praxis bedeutete dies, dass es als legitim angesehen werden würde, die Wertigkeit einer Person oder eines Objekts auf der Grundlage seines Rufs festzulegen. Im Fall von Schule zeigt sich diese Werteordnung beim Schulwahlverhalten von Eltern. So ist für viele der Ruf einer Schule entscheidend dafür, ob sie ihr Kind an einer Schule anmelden oder nicht.

Welt der Innovation: „Diese Konvention begründet die Wertigkeit von Personen, denen besondere kreative Fähigkeiten und Leistungen zuerkannt werden“ (ebd., S. 145), bzw. kann sich auch die Wertigkeit von Objekten mit dem Grad der Innovation begründen lassen. Diese Rechtfertigungsordnung wird von Diaz-Bone (2015) als unsicher und unberechenbar beschrieben, weil die Frage danach, ob etwas inspirierend und innovativ ist, nicht objektiv zu beantworten ist.

Welt des Staatsbürgers: „Diese Konvention hält eine Welt im Inneren zusammen, die die Solidarität und die Rechte und die Gleichheit der Menschen betont, weil sie Teil eines aufgeklärten, bürgerlichen Kollektivs sind (Parteien, Staat, Vereine, soziale Bewegungen, Initiativen, etc.)“ (ebd., S. 146). „In Organisationen gibt es entsprechende Abteilungen oder Positionen (Gleichstellungsbeauftragte), die gegen Diskriminierung (hinsichtlich Geschlecht und ethnischer Zugehörigkeit) vorgehen und auf die Gleichbehandlung achten sollen“ (ebd., S. 147).

Die einzelnen Werteordnungen unterscheiden sich unter anderem darin,

- **Worin sich die Wertigkeit äußert**, z. B. bei der „Welt der Innovation“ ist es der Grad der Kreativität;
- **Was relevante Informationen zur Messung sind**, z. B. ist es häufig der monetäre Preis bei der „Welt des Marktes“;
- **Wovon es abhängt, dass Wert zugeschrieben wird**, z. B. bei der „Welt des Hauses“ ist es der Grad an Vertrauen;
- **Wie das Objekt wahrgenommen wird**, z. B. bei der „Welt der Meinung“ ist das relevante Attribut eines Objekts die Bekanntheit (vgl. Boltanski und Thévenot 1999, S. 368).

Die oben beschriebenen Werteordnungen wurden aus sozialphilosophischen Schriften abgeleitet und empirisch durch diverse Untersuchungen gestützt (vgl. Boltanski und Thévenot 2007). Dabei argumentieren Boltanski und Thévenot, dass die einzelnen Rechtfertigungsordnungen systemübergreifend auftauchen können und nicht nur für einen Bereich gelten (vgl. Beckert und Musselin 2013, S. 12), z. B. die ökonomische Argumentation („Welt des Marktes“) beschränkt sich nicht auf das Wirtschaftssystem, bzw. werden im Wirtschaftssystem auch andere Werteordnungen für legitim erachtet (vgl. Diaz-Bone 2015, S. 141).

Grundsätzlich sind die Werteordnungen als gleichwertig anzusehen (vgl. Stark 2009, S. 24). So gibt es keine Letztbegründung, die erklärt, warum bestimmte Werteordnungen höher anzusiedeln sind als andere. Stark (2009) spricht in diesem Zusammenhang von einer Heterarchie von Werteordnungen. In Bezug auf konkrete Handlungsprozesse ist es jedoch notwendig,

eine Entscheidung darüber zu treffen, welche Werteordnungen angewandt bzw. miteinander kombiniert werden sollen.

In Bezug auf diese Arbeit lassen sich die theoretischen Überlegungen zu Rechtfertigungs- bzw. Werteordnungen auf verschiedene Aspekte anwenden: Beispielsweise wenn begründet werden soll, welches Verständnis von guter Schule in Form von Indikatoren operationalisiert werden sollte. So könnte argumentiert werden, dass eine Schule gut ist, wenn sie im Wohnort eine hohe Reputation genießt („Welt der Meinung“) oder sich stets weiterentwickelt („Welt der Innovation“). Ausgehend von der Fragestellung der Arbeit dient das theoretische Konzept der Rechtfertigungsordnungen im Folgenden primär als Grundlage, um zu verstehen, **wie die Mitarbeiter_innen von Schulinspektionen bei der Entwicklung von Indikatoren zu Inklusion vorgegangen sind bzw. wie sie ihr Handeln begründen**. Es geht daher um die Frage, wie begründet wird, welche Rahmenbedingungen bei der Entwicklung von Indikatoren für relevant erachtet wurden.

4.3 Exkurs: Qualität von Schule

Die Instrumente zur Messung von Inklusion repräsentieren gemeinsam mit den anderen Messinstrumenten der jeweiligen Schulinspektionen bestimmte Verständnisse von guter Schule. Ausgehend von den vorangegangenen Überlegungen stellt sich die Frage, inwiefern Indikatoren zur Messung von Schulqualität bereits etabliert sind und somit nicht mehr in Frage gestellt werden. Dabei bezieht sich diese Frage nicht nur auf Schulinspektionen, sondern generell auf den Diskurs zu Schulqualität.

Ob es im deutschsprachigen Diskurs Konsens darüber gibt, was unter Schulqualität zu verstehen ist, ist auch in Bezug auf die Entwicklung von Indikatoren zur Umsetzung von Inklusion eine relevante Frage: Bei einem dominanten Verständnis von guter Schule liegt es nahe, Inklusion so zu deuten, dass es kohärent zu diesem Verständnis ist.

„Von der Qualität der Schule wird sowohl in der bildungsinteressierten Öffentlichkeit als auch in schulpädagogischen Fachdiskussionen so geredet, als ob es überhaupt keinen Zweifel darüber geben könnte, was damit gemeint ist“ (Heid 2007, S. 55).

Wie bereits erwähnt gibt es jedoch „keineswegs Konsens darüber (...), was genau unter Qualität zu verstehen ist“ (Ditton 2000, S. 74). So gibt es unterschiedliche Vorstellungen von Qualität, die unkoordiniert verwendet werden (vgl. Hupka-Brunner et al. 2015, S. 8). Demnach fehlt es häufig an einer Diskussion darüber, anhand welcher Vorstellungen etwas für gut befunden wird oder nicht. Beispielsweise versteht der Bildungsjournalist Christian Füller unter „guter“ Schule Folgendes: „Gute Schule ist eine Schule vor Ort, die phantastische Kinder zur Entfaltung ihrer nahezu unbegrenzten Möglichkeiten anregt“ (Füller 2009, S. 177). Dagegen gibt es auch andere Verständnisse von „guter“ Schule, die sich beispielsweise stärker an dem Erreichen von Bildungsstandards orientieren, z. B. wenn die Ergebnisse der IQB-Vergleichsarbeiten dazu genutzt werden, um Aussagen über die Qualität von Einzelschulen zu treffen. Häufig unterscheiden sich die Vorstellungen von „guter“ Schule darin, was als primärer Zweck von Schule angesehen wird und wie diese Funktionen erreicht werden sollen. In diesem Zusammenhang bietet es sich an, erneut auf die gesellschaftlichen Funktionen von Schule (→ Kapitel 2) zurückzugreifen: So unterscheidet Fend (2008) zwischen vier Funktionen von Schule: Qualifikations-, Selektions-, Integrations- und Enkulturationsfunktion (nähere Ausführung → Kapitel 2). Welche Relevanz den einzelnen Funktionen beigemessen wird und wie diese inhaltlich gefüllt werden, kann

variieren. Dies spiegelt sich in unterschiedlichen Vorstellungen von „guter“ Schule wider, die wiederum auf unterschiedlichen Werteordnungen basieren.

Ditton und Müller benennen zwei Verständnisse von Schulqualität, die ihres Erachtens den deutschsprachigen Diskurs dominieren (vgl. Ditton und A. Müller 2011, S. 101). So gibt es zum einen die so genannte „funktionalistische Perspektive“ auf Schule und zum anderen eine „humanistische Perspektive“ (vgl. ebd., S. 101).

„Funktionalistische“ Perspektive auf Schulqualität Dieses Qualitätsverständnis orientiert sich primär an der Qualifikationsfunktion, wonach Schule die Vermittlung von Kompetenzen zur gesellschaftlichen Teilhabe ermöglichen soll (vgl. Fend 2006). Unter Qualifikation durch Schule wird verstanden, dass Schüler_innen Kompetenzen erwerben, die ökonomisch verwertbar sind, z. B. den Zugewinn an kognitiven Kompetenzen (vgl. Ribolits 2009, S. 15). Die Qualität einer Einzelschule bemisst sich daran, inwiefern die Schüler_innen bestimmte Kompetenzen erwerben. In Bezug auf Messinstrumente spielt sich daher jenes Verständnis in Schulleistungsstudien wie PISA, IGLU oder TIMSS wider. Dort werden beispielsweise die mathematischen und naturwissenschaftlichen Fähigkeiten von Schüler_innen erfasst. Das oben beschriebene Qualitätsverständnis steht jedoch aus diversen Gründen in der Kritik: Dies liegt unter anderem daran, dass der Fokus primär darauf gelegt wird, welche Leistungen die Schülerinnen „zu einem bestimmten Zeitpunkt (...) in bestimmten Schulfächern oder erworbenen Basiskompetenzen in bestimmten Domänen“ (Scharenberg 2015, S. 54) erreicht haben. Holtappels weist daraufhin, dass neben den kognitiven Kompetenzen in die Schule einen Beitrag dazu leisten soll, dass die Schüler_innen noch weitere Fähigkeiten erwerben können: z. B.

„überfachliche Schlüsselkompetenzen und metakognitive Fähigkeiten (z. B. Analyse- und Planungskompetenzen, kommunikative Fähigkeiten) (...). Daneben gehören aber ebenso die Entwicklung sozialer Kompetenzen und Ich-Kompetenzen im Sinne einer Unterstützung der Persönlichkeitsbildung und Identitätsentwicklung der Schüler/-innen“ (Holtappels 2003, S. 37).

Des Weiteren lassen sich aus der Messung der Outputs keine Aussagen darüber treffen, welche Rahmenbedingungen und Prozesse zu den jeweiligen Ergebnissen führen (vgl. ebd., S. 39). Gleichwohl lässt jenes Verständnis von Schulqualität bzw. die Operationalisierung Aussagen darüber zu, inwiefern die Schüler_innen zu einem bestimmten Zeitpunkt über die abgefragten Kompetenzen verfügen und ob es Unterschiede zwischen einzelnen Gruppen von Schüler_innen gibt.

In Bezug auf den Gegenstand der Arbeit stellt sich die Frage, ob jenes Verständnis von guter Schule mit dem Verständnis von Inklusion, das im Rahmen der Arbeit als normative Vergleichsfolie dient, kohärent ist: Wie bereits in den theoretischen Überlegungen zu Inklusion ausgeführt (→ Abschnitt 2.2.4), widersprechen sich jene Verständnisse in der Theorie nicht, können jedoch in der Praxis bestimmte Handlungen implizieren, die zu konträren Wirkungen führen, z. B. Selektionsverhalten bei der Aufnahme von Schüler_innen.

Humanistische Perspektive auf Schulqualität Bei diesem Qualitätsverständnis steht „die allgemeine Entwicklung aller menschlichen Kräfte“ (Ditton und A. Müller 2011, S. 101) im Vordergrund. Im Gegensatz zum vorherigen Verständnis steht dabei nicht die ökonomische Verwertbarkeit, sondern die Persönlichkeitsentwicklung der Schüler_innen im Mittelpunkt. Dies kann den Erwerb von kognitiven Kompetenzen einschließen, aber auch andere Domänen,

z. B. Identitätsbildung, soziale Kompetenzen, usw. werden in den Blick genommen. Hierbei wird an reformpädagogische Ansätze angeknüpft, wonach Lernen „als Entwicklung im Kinde aus und als Lernen vom Kinde aus verstanden wird“ (Fend 2000, S. 57). Dabei liegt bei diesem Qualitätsverständnis der Fokus nicht auf dem zu erreichenden Output in Bezug auf das einzelne Kind, sondern auf den schulischen Prozessen. So soll „Schule für möglichst alle Schüler zu produktiven Räumen des Lernens und ihrer langfristigen Entwicklung werden“ (ebd., S. 57). In Bezug auf das in dieser Arbeit verwendete Verständnis von Inklusion lassen sich hier insofern Anknüpfungspunkte finden, als beispielsweise das Erfahren von intersubjektiver Anerkennung für die Persönlichkeitsentwicklung maßgeblich ist (nähere Ausführung → Kapitel 2).

Abschließend lässt sich festhalten, dass es im deutschsprachigen Diskurs keinen Konsens darüber gibt, was unter einer guten Schule verstanden wird. Generell lässt sich festhalten, dass es bislang an einer empirischen Analyse darüber fehlt, welche Verständnisse von guter Schule den deutschsprachigen Diskurs prägen und worin sie sich unterscheiden. Für das Handeln der Schulinspektion hat dies zur Folge, dass sie nicht auf ein bereits etabliertes Set an Indikatoren zur Messung von Schulqualität zurückgreifen kann. Es ist daher davon auszugehen, dass es in den Schulinspektionen ein Aushandlungsprozess darüber stattgefunden hat, welche Verständnisse von guter Schule durch die Indikatoren repräsentiert werden sollen. Das Vorhandensein von unterschiedlichen Verständnissen kann zur Folge haben, dass diese bei der Entwicklung von Indikatoren miteinander kombiniert werden und somit Messinstrumente der Schulinspektionen verschiedene Verständnisse von guter Schule repräsentieren. Dies ist insofern nicht überraschend, als sich die Verständnisse von guter Schule unter anderem darin unterscheiden, welche Erwartungen an Schule formuliert werden. Eine Kombination liegt daher nahe, da an Schule verschiedene Erwartungen formuliert werden können und sich diese nicht immer widersprechen.

In Bezug auf die unterschiedlichen Verständnisse von Schulqualität stellt sich darüber hinaus die Frage, woher die Vorstellungen von „guter“ Schule kommen. Fend weist darauf hin, dass „unsere Vorstellungen von guter Schule mit unseren Bildern darüber verwoben sind, wie menschliches Zusammenleben insgesamt aussehen sollte“ (Fend 1994, S. 24; siehe auch Terhart 2000, S. 816). Als weitere Faktoren nennt Terhart „Traditionen, allgemein- und sozialpolitische Überzeugungen, sozial- und bildungstheoretische Erwägungen und Entscheidungen“ (ebd., S. 816). So spiegeln sich in Vorstellungen von guter Schule beispielsweise unterschiedliche Konzepte von Gerechtigkeit, variierende Menschenbilder sowie verschiedene Leistungsbegriffe wider. Des Weiteren spielt eine Rolle, welche Überzeugungen zur Bildungsfähigkeit vorhanden sind, z. B. „naturalistischer Glaube an Bildung“ vs. „universelle Bildungsfähigkeit“ (Gomolla 2009, 32f.). In Bezug auf diese Arbeit ist bei der empirischen Erhebung zu berücksichtigen, welche Verständnisse von guter Schule bei den Entwickler_innen der Instrumente vorhanden sind und inwiefern diese möglicherweise prägend wirkten, wenn es um die Entwicklung von Indikatoren zu Inklusion ging.

4.4 Zwischenfazit

Durch die bewertungssoziologische Perspektive auf den Gegenstand der Arbeit sind verschiedene Aspekte deutlich geworden: Indikatoren stellen soziale Konstruktionen dar. Durch die Entscheidung für konkrete Indikatoren und gegen alternative Varianten werden bestimmte Perspektiven auf die Wirklichkeit abgebildet und andere werden vernachlässigt. Des Weiteren gehen bestimmte Implikationen damit einher, wenn Indikatoren dafür genutzt werden

numerische Werte zu erheben, z. B. Fokus auf die Erhebung von Informationen, die messbar sind. In Bezug auf die Entwicklung von Indikatoren lässt sich festhalten, dass es kaum Studien gibt, die sich mit der Entwicklung von Indikatoren auseinandersetzen. Erste Hinweise zur Analyse des empirischen Materials ergeben sich möglicherweise aus den jeweiligen Bewertungskonstellationen sowie aus der Notwendigkeit, dass Indikatoren einen bestimmten Grad an Institutionalisierung benötigen, um von Relevanz zu sein.

Durch den Exkurs zur Qualität von Schule ist deutlich geworden, dass es keinen Konsens darüber gibt, was eine gute Schule kennzeichnet bzw. wie dies gemessen wird. Dies trifft auch auf Indikatoren zu Inklusion zu (nähere Ausführung → Kapitel 2), die einen Teilbereich des gesamten Instrumentensets von Schulinspektionen darstellen können. Somit kann auch diesen Indikatoren die Funktion zugeschrieben werden, dass sie bestimmte Verständnisse von guter Schule repräsentieren. Aufgrund der fehlenden Existenz von etablierten Indikatorensets zu Inklusion und zu guter Schule kann es zu den oben beschriebenen „critical moments“ kommen. Dabei wird zum einen auf Rechtfertigungsordnungen zurückgegriffen, um zu begründen, welches Verständnis von guter Schule von Relevanz ist, und zum anderen wird auf Rechtfertigungsordnungen zurückgegriffen, um das Vorgehen bei der Entwicklung von Indikatoren zu begründen. Letzteres wird im Rahmen der Arbeit empirisch untersucht. Des Weiteren wurde in Bezug auf Schulinspektionen deutlich, dass diese im Vergleich zu anderen Evaluationsverfahren, die von Organisationen wie der Bildungsadministration, privaten Agenturen oder der Bildungsforschung realisiert werden, abweicht: Erstens wird nicht nur auf quantitative Indikatoren zurückgegriffen, sondern es werden auch Formen der qualitativen Datenerhebung angewandt. Auch wenn es möglich wäre in bestimmten Teilbereichen Vergleiche zwischen Schulen herzustellen, findet dies in der Praxis nicht statt. Demnach werden auch keine Rankings veröffentlicht, die Aussagen darüber zulassen, welche Schule die „beste“ oder die „schlechteste“ im jeweiligen Bundesland ist.

5 Sensemaking-Ansatz

Im Rahmen dieser Arbeit dient der Sensemaking-Ansatz als theoretische Brille, um zu analysieren, wie die Instrumenteentwickler_innen bei der Deutung von Inklusion vorgegangen sind, wie die Messinstrumente zu Inklusion entstanden sind und welche Rahmenbedingungen die Deutungsprozesse prägten.

Generell wird der Sensemaking-Ansatz verwendet, um Entscheidungsprozesse in Organisationen oder (häufig auch) Unfallszenarien zu erklären (für einen Überblick siehe Sandberg und Tsoukas 2015). In Bezug auf den Bildungsbereich sind insbesondere die Arbeiten von Spillane und Coburn zu nennen, die sich mit der Implementation von politischen Vorgaben im schulischen Kontext auseinandersetzen (z. B. Spillane 2004; Coburn 2016; für einen Überblick siehe DeMatthews 2012; Quinn 2009). Während es diverse Studien zu Deutungsprozessen von Schulleitungen gibt, lässt sich bezogen auf den Gegenstand dieser Arbeit – Messinstrumente der Schulinspektionen in Hinblick auf die Umsetzung von Inklusion – festhalten, dass es kaum empirische Untersuchungen gibt, die Anknüpfungspunkte bieten. So ist der Autorin lediglich die Studie von DeMatthews (2012) bekannt, der mithilfe des Sensemaking-Ansatzes analysiert, wie Inklusion von verschiedenen Schulleitungen eines US-amerikanischen Bezirks gedeutet wird (nähere Ausführung des Forschungsstands in Bezug auf die Implementation von politischen Vorgaben → Abschnitt 5.2). Im Zusammenhang mit Evaluationen ist noch die Studie von Halverson et al. (2004) zu nennen, die sich mit dem Sensemaking von Schulleitungen auseinandersetzen und dabei in den Blick nehmen, wie Lehrkräfte von diesen evaluiert werden.

Im Folgenden wird nun zunächst auf die theoretischen Überlegungen von Weick (1995) eingegangen. Dieser hat den Sensemaking-Ansatz entwickelt, um Deutungen von Ereignissen oder Situationen zu analysieren. Anschließend folgt eine Darstellung, wie sich der Ansatz auf die Implementation von politischen Vorgaben anwenden lässt. So haben Spillane et al. (2002) den Ansatz von Weick hierfür adaptiert und konkretisiert.

5.1 Sensemaking-Ansatz nach Weick

Um das Potential des Sensemaking-Ansatzes zu beschreiben, wird in der Literatur häufig auf Weicks Analyse des „Mann Gulch Waldbrand“ (vgl. Weick 1993) verwiesen: Im August 1949 brach in der Mann Gulch Gegend in Montana ein riesiges Feuer aus, welches viele Todesopfer unter den Feuerwehrmännern forderte. In seiner Analyse von Interviews und anderen Formen der Dokumentation beschäftigt sich Weick mit der Frage, wie die einzelnen Feuerwehrmänner die Situation deuteten. Hierbei arbeitet er verschiedene Ursachen für Fehldeutungen heraus, z. B. fehlende Erfahrung der Feuerwehrmänner, Panikanfälle, Missinterpretationen der Situation oder fehlendes Vertrauen gegenüber erfahreneren Kollegen. In seiner Analyse nimmt Weick demnach nicht nur in den Blick, wie die Situation von Seiten der Feuerwehrmänner gedeutet wurde, sondern auch, auf welche Wissensbestände hierbei zurückgegriffen wurde und welche Handlungen daraus folgten.

Dabei argumentiert Weick, dass Sensemaker beim Deuten einer Situation, auf einen bereits vorhandenen Referenzrahmen zurückgreifen. Dieser ermöglicht es Informationen aufzunehmen, zu verarbeiten und zu nutzen (Westley 1990 zit. In Weick 1995, S. 4). Dabei gehören zum

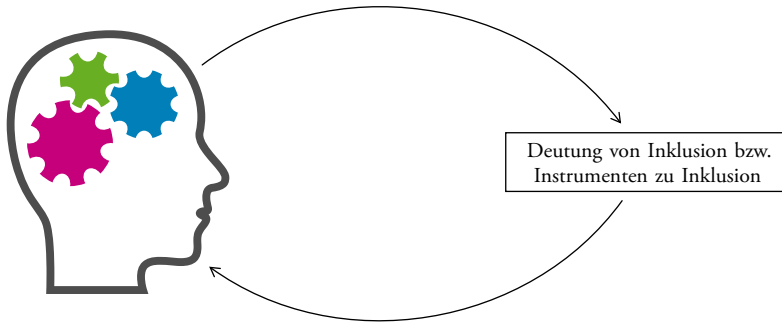


Abb. 5.1: Sensemaking nach Weick

Referenzrahmen unter anderem Erwartungen, Wissensbestände, Routinen, Vorannahmen oder akzeptierte Rechtfertigungen (vgl. Weick et al. 2005, S. 409). In dem Fall der Feuerwehrmänner gehören zum Referenzrahmen beispielsweise bisherige Erfahrungen mit ähnlichen Ereignissen. In Bezug auf das Verhältnis zwischen Referenzrahmen und der zu deutenden Situation ist festzuhalten, dass die Beziehung stets zirkulär ist (siehe Abb. 5.1), d. h. im Rahmen des Deutungsprozesses kann sich auch der bestehende Referenzrahmen verändern bzw. sich anpassen (vgl. Spillane 2004, S. 76).

Neben der Analyse von tragischen Unfällen bietet es sich an, auch andere Situationen zu analysieren. So finden das Deuten von Situationen und sich daraus ergebende Handlungen ständig statt. Nach Weick liegt es jedoch nahe, Ereignisse oder Prozesse zu betrachten, die durch Mehrdeutigkeit oder Ungewissheit gekennzeichnet sind (vgl. Weick 1995, S. 91): Im Fall von Mehrdeutigkeit realisieren Menschen Sensemaking, weil sie sich für eine von mehreren Interpretationen entscheiden müssen (vgl. ebd., S. 91). Im Fall von Ungewissheit ist es so, dass ihnen unbekannt ist, ob es vorhandene Deutungen gibt (vgl. ebd., S. 91). Der weitere Verlauf des Deutungsprozesses ist abhängig von dem jeweiligen Merkmal: Im Fall von Mehrdeutigkeit stehen genug Informationen zur Verfügung, „vielmehr bedarf es einer anderen <Qualität>“ (Appelt 2016, S. 57). Demnach erfordert es die Situation, dass sich der Sensemaker für eine bestimmte Deutung entscheidet. Dagegen ist es im Fall von Ungewissheit notwendig, mehr Wissen zu erlangen, um etwas deuten zu können. Im Rahmen von empirischen Untersuchungen bietet sich die Analyse von mehrdeutigen oder ungewissen Situationen an, weil in diesen Fällen zum einen die Deutungsprozesse komplex sind und häufig Erklärungsbedarf entsteht, wie es zu den jeweiligen Deutungen gekommen ist. Demnach liegt die Anwendung des Ansatzes bei der Deutung von Inklusion bzw. der Entwicklung von Messinstrumenten zu Inklusion nahe: So treffen beide von Weick genannten Merkmale in diesem Fall zu: Es gibt verschiedene Deutungen von Inklusion (~ Mehrdeutigkeit) und es gibt keine Messinstrumente zur Umsetzung von Inklusion (~ Ungewissheit). Mitarbeiter_innen der Schulinspektion stehen daher vor der Herausforderung, Inklusion als Konzept zu deuten. Unter der Annahme, dass es Konsens darüber geben würde, wie Inklusion zu definieren und deren Umsetzung empirisch zu messen ist, wäre es irrelevant zu untersuchen, wie Inklusion auf der Ebene der Schulinspektionen rekontextualisiert wurde. In diesem Fall hätten die Instrumenteentwickler_innen kaum Handlungsspielraum für Deutungen und es gebe zwischen den Bundesländern keine Unterschiede.

5.1.1 Kennzeichen von Sensemaking nach Weick

Weick (1995) führt sieben Kennzeichen von Sensemaking auf, die den Deutungsprozess charakterisieren. Diese werden im Folgenden näher beschrieben. Hierbei wird auch auf bestehende Kritikpunkte an Weicks Arbeiten eingegangen bzw. werden Herausforderungen, die ins Besondere im Zusammenhang mit dem empirischen Design entstehen können, thematisiert.

„Grounded in Identity“: Sensemaking ist geprägt durch die Auseinandersetzung des Sensemakers mit seiner Umwelt. Wie der Sensemaker seine Umwelt wahrnimmt, beziehungsweise was als relevante Umwelt gesehen wird, ist abhängig davon, wie dieser sich selbst definiert (vgl. ebd., S. 20). Zugleich prägt auch die Umwelt die Identität des jeweiligen Sensemakers (vgl. ebd., S. 20). Unter Bezugnahme auf Erez & Earley (1993 zit. In Weick 1995) argumentiert Weick, dass die Entwicklung beziehungsweise das Aufrechterhalten einer Identität durch drei Bedürfnisse geprägt sei: Streben nach einem positiven Selbstbild, das Bedürfnis nach Selbstwirksamkeit und das eigene Selbstbild muss in sich stimmig und widerspruchsfrei sein (vgl. ebd., S. 20). Weick argumentiert, dass sich diese Bedürfnisse auch beim Sensemaking zeigen. Darüber hinaus ist zu berücksichtigen, dass bei Deutungsprozessen in Organisationen nicht nur eine Rolle spielt, wie der Sensemaker sich als Individuum, sondern auch, wie er sich als Repräsentant der jeweiligen Organisation versteht (vgl. ebd., S. 23). Es geht daher nach DeMatthews nicht nur um die Frage „Who am I?“, sondern auf der Organisationsebene auch um die Frage „Who are we?“ (DeMatthews 2012, S. 56).

„Retrospective“: Weick versteht die Wirklichkeit als einen kontinuierlichen Strom an Erfahrungen (vgl. Weick 1995, S. 25). Für Deutungsprozesse muss dieser kontinuierliche Strom verlassen werden, da Erfahrungen erst dann als einzelne Episoden greifbar bzw. gedanklich bearbeitbar werden. Sensemaking findet demnach statt, wenn etwas bereits geschehen ist (vgl. ebd., S. 26). Hierbei ist jedoch zu berücksichtigen, dass das Sensemaking zeitlich sehr nah an die erlebte Situation anknüpft. Beispielsweise deuten im Fall des oben beschriebenen Mann Gulch Waldbrands die Feuerwehrmänner die Situation, während sich das Feuer weiter ausbreitet. Das Sensemaking ist daher notwendig, um handlungsfähig zu bleiben.

Im Hinblick auf den temporalen Charakter des Sensemakings wird Weick häufig dafür kritisiert, dass er die prospektive Perspektive vernachlässigen würde (für einen Überblick siehe Sandberg und Tsoukas 2015, S. 519; siehe auch Nijhof und Jeurissen 2006, S. 321). Sandberg und Tsoukas (2015) argumentieren jedoch, dass auch auf die Zukunft bezogene Denkprozesse beim Sensemaking retrospektiver Natur wären, da im Denken antizipiert wird, welche Implikationen sich aus den jeweiligen Deutungen ergeben würden (vgl. ebd., S. 519). Demnach wäre dies kein Widerspruch zu Weicks Arbeiten und die prospektive Perspektive wäre nach Weick ein Denken im „future perfect tense“ (Weick 1979) und wird somit unter dem Merkmal „retrospective“ berücksichtigt.

In Bezug auf empirische Untersuchungen ist bei der Entwicklung des Forschungsdesigns und der Auswertung der Daten kritisch zu reflektieren, wie viel Zeit zwischen der zu deutenden Situation und dem Deuten der Situation liegt. So besteht die Gefahr, dass im Rahmen von empirischen Erhebungen nicht das tatsächliche Sensemaking, sondern ein Sensemaking vom Sensemaking abgebildet wird. Dies ist beispielsweise der Fall, wenn in Interviews die Erzählungen von den tatsächlich stattgefundenen Deutungsprozessen abweichen, weil diese dem Sensemaker nicht mehr genau in Erinnerung sind bzw. sozial erwünscht geantwortet wird (nähere Ausführung → Kapitel 8).

- „**Enactive of sensible environments**“: Sensemaker schaffen sich ihre Umwelt selbst. Dabei entscheidet der Sensemaker, was als Umwelt wahrgenommen und was ignoriert wird. Durch ihr eigenes Handeln gestalten Menschen für sich eine Umgebung, die zu einem Hindernis, aber auch zu einer Möglichkeit werden kann (vgl. Weick 1995, S. 31). Hierbei ist zu beachten, dass Menschen in ihrem Alltag häufig verdrängen, dass die von ihnen wahrgenommene Realität stets von ihnen selbst konstruiert wird und nehmen diese als naturgegeben an (ebd., S. 23). Dieses Verhalten kann als Schutzmechanismus gelesen werden: So würde das ständige Hinterfragen der eigenen Wahrnehmung dazu führen, dass Menschen nicht mehr handlungsfähig sind.
- „**Social**“: Der soziale Kontext spielt beim Sensemaking-Ansatz eine zentrale Rolle. Während des Deutens greifen Sensemaker auf Verständnisse und Normen zurück, die sie im Rahmen ihrer Sozialisation erworben haben. Des Weiteren beziehen sie den sozialen Kontext, in dem sie sich aktuell befinden, bei konkreten Deutungsprozessen mit ein. In diesem Zusammenhang wird beim Sensemaking häufig eine (fiktive) Audienz antizipiert. So ergibt sich für die Sensemaker die Notwendigkeit, das eigene Handeln gegenüber dieser Audienz legitimieren zu können (vgl. ebd., S. 40). Demnach ist auch der soziale Kontext ein Teil des Referenzrahmens.
- „**Ongoing**“: „Sensemaking never starts. The reason it never starts is that pure duration never stops“ (ebd., S. 43). Demnach finden im Alltag ständig Deutungen von Situationen statt. Hierbei wird meistens auf bestehende Deutungsmuster zurückgegriffen und diese werden reproduziert. Gleichwohl besteht auch die Möglichkeit, dass sich Deutungen im Laufe der Zeit verändern. In Bezug auf empirische Untersuchungen impliziert jenes Kennzeichen, dass stets nur eine Momentaufnahme bzw. im Fall von Längsschnittstudien mehrere Momentaufnahmen gezeichnet werden können. Angewandt auf den Gegenstand dieser Arbeit bedeutet dies: Es ist möglich, dass sich die Deutungen von Inklusion bei den Instrumenteentwickler_innen im Laufe der Zeit verändern. Die im Rahmen der Arbeit ausgewählte Momentaufnahme ist jedoch insofern von Relevanz, da die Deutungen der Instrumenteentwickler_innen maßgeblich die Entwicklung der Messinstrumente zur Umsetzung von Inklusion prägten. Letzere weisen durch ihre Standardisierung und Anwendung in der Praxis einen gewissen Grad an Institutionalisierung auf.
- „**Focused on and by extracted cues**“: Sensemaking ist das Deuten von ungewissen oder mehrdeutigen Situationen. Hierbei werden sogenannte „cues“ konstruiert (vgl. ebd., S. 55), d. h. einzelne Aspekte einer Situation werden in den Blick genommen, um die Situation als solche gedanklich bearbeitbar zu machen. Diese Indikatoren geben Hinweise darauf, wie eine Situation wahrgenommen werden kann und auf welcher Grundlage Entscheidungen getroffen werden. Bezüglich abstrakter Konzepte wie Inklusion kann beispielsweise die Zielgruppe von Inklusion ein möglicher „cue“ sein. Was als „cue“ ausgewählt und wie dieser interpretiert wird, ist abhängig vom jeweiligen Kontext (vgl. ebd., S. 51) bzw. vom Sensemaker. In diesem Zusammenhang bietet es sich in der empirischen Analyse an, nicht nur die ausgewählten „cues“ zu beachten, sondern auch, welche Aspekte von Seiten der Instrumenteentwickler_innen nicht beachtet wurden.
- „**Driven by plausibility rather than accuracy**“: Nach Weick spielt es für den Sensemaker weniger eine Rolle, ob das Vorgehen bei Deutungsprozessen sorgfältig bzw. akkurat ist. Dies wäre der Fall, wenn neue Deutungen als Arbeitshypothesen gesehen werden würden und es stets derer Verifizierung bedürfte. Stattdessen ist für die Sensemaker von Relevanz, dass die

eigene Deutung unter Berücksichtigung der relevanten Umwelt für plausibel gehalten wird (vgl. ebd., 55f.). Demnach geht es darum, ein kohärentes Bild zu erzeugen.

In Anbetracht der sieben Kennzeichen nach Weick (1995) lässt sich Folgendes festhalten: Der Sensemaking-Ansatz legt den Fokus auf die Interaktion einer zu deutenden Situation und dem bestehenden Referenzrahmen (→ Abb. 5.1). Dieser Prozess kann aus mehreren Gründen als sehr dynamisch beschrieben werden: **Erstens**, was als Referenzrahmen gesehen wird, kann in Abhängigkeit vom Sensemaker, dem Setting und der jeweiligen Situation variieren. **Zweitens**, aus der Interaktion zwischen der zu deutenden Situation und dem Referenzrahmen können sich unterschiedliche Implikationen ergeben, z. B. Anpassung des Referenzrahmens. **Drittens**, Sensemaking ist ein fortlaufender Prozess, der nie abgeschlossen ist.

5.1.2 Sensemaking in Organisationen

Bislang lag der Fokus primär auf Deutungsprozessen auf der individuellen Ebene. Im Rahmen dieser Arbeit geht es jedoch nicht nur um Deutungsprozesse von Einzelpersonen. Es geht auch darum, wie innerhalb der Schulinspektionen in den einzelnen Bundesländern ein Austausch stattgefunden hat und welche Deutungen von Inklusion sich in den Messinstrumenten der Schulinspektionen wiederfinden. Im Folgenden wird näher auf Weicks Verständnis von Organisationen bzw. Prozessen des Organisierens eingegangen, um das Verhältnis zwischen den einzelnen Ebenen zu beschreiben.

Weick versteht unter Prozessen des Organisierens „fortlaufende unabhängige Handlungen zu vernünftigen Folgen zusammenzufügen, sodass vernünftige Ergebnisse erzielt werden“ (Weick 1985, S. 11). Hierbei geht es darum, Mehrdeutigkeit zu reduzieren und die „Spannweite der Möglichkeiten zu verkleinern“ (ebd., S. 15). Um dies zu erreichen, versuchen die Personen, die in den Prozess des Organisierens involviert sind, durch Selektion und Interpretation bestimmte Bedeutungen zu stabilisieren (vgl. ebd., S. 16). Wenn mehrere Personen involviert sind, die wechselseitig in ihren Handlungen voneinander abhängig sind, wird dieser Prozess als Organisieren bezeichnet (ebd., S. 16). Dieses Verständnis von Organisationen ist für Weick stark mit dem Konzept des Sensemaking verknüpft: „The operative image of organization is one in which organization emerges through sensemaking, not one in which organization precedes sensemaking or one in which sensemaking is produced by organization“ (Weick et al. 2005, S. 410). Dabei ist ein maßgebliches Merkmal des Organisierens die Kommunikation zwischen Akteuren. So ist der sprachliche Austausch grundlegend, um auf kollektiver Ebene deuten zu können (vgl. ebd., S. 413).

Mit Bezug auf Wiley (1988 zit. In Weick 1995, S. 70) unterscheidet Weick zwischen verschiedenen Ebenen des Sensemakings: der individuellen, intersubjektiven, generisch-subjektiven und der extrasubjektiven Ebene. Als intersubjektive Ebene wird der Austausch über Bedeutungen zwischen mindestens zwei Personen verstanden (vgl. ebd., S. 70). Auf der generisch-subjektiven Ebene verlieren konkrete Personen an Bedeutung und stattdessen sind soziale Strukturen dominant (vgl. ebd., S. 70). Die extrasubjektive Ebene befindet sich auf einer noch abstrakteren Ebene, die keinem konkreten Subjekt bzw. Objekt zugeschrieben werden kann. Als Beispiel für diese Ebene wird der Kapitalismus oder die Mathematik angeführt (ebd., S. 72). Demnach sind es tendenziell eher Heuristiken, die unter dieser Ebene gefasst werden.

Prozesse des Organisierens sind nach Weick zwischen der intersubjektiven und der generisch-subjektiven Ebene zu verorten. So finden zwischen Individuen Absprachen über Deutungen statt (intersubjektiv), die sich wiederum in Routinen und Strukturen widerspiegeln können

(subjektiv-generische Ebene). Angewendet auf den Gegenstand dieser Arbeit bedeutet dies, dass im empirischen Teil verschiedene Deutungsprozesse in den Blick genommen werden: die individuelle Ebene durch die Deutungen der Instrumenteentwickler_innen, der Austausch zwischen den Instrumenteentwickler_innen einer Schulinspektion auf intersubjektiver Ebene sowie die Messinstrumente, die der subjektiv-generischen Ebene zugeordnet werden können.

In Bezug auf die intersubjektive Ebene ist zu ergänzen, dass es bislang wenig Wissen darüber gibt, welche Prozesse beim kollektiven Sensemaking stattfinden (vgl. Merkus et al. 2016). Maitlis und Christianson vergleichen die Prozesse mit einem Jazzorchester: „Members must listen closely to each other take turns leading and following, and respond together in real-time to novel or unexpected performance“ (Maitlis und Christianson 2014, S. 78). Im Rahmen dieser Arbeit wird die intersubjektive Ebene, d. h. die gemeinsame Deutung von Inklusion innerhalb einer Schulinspektion, betrachtet, jedoch sind hierzu — aufgrund des empirischen Designs — nur begrenzt Aussagen möglich (nähere Ausführung → Kapitel 6).

Im Hinblick auf die einzelnen Ebenen weist Wetzel darüber hinaus auf das Spannungsfeld zwischen Innovation und Routine hin: „Während auf intersubjektiver Ebene Neues durch konkrete persönliche Beziehungen geschaffen, aufgebaut und implementiert werden kann, kontrollieren Organisationen als generisch-subjektive Formen genau diese innovierende Intimität“ (vgl. Wetzel 2001, S. 178). Demnach zeigen sich Routinen auf der generisch-subjektiven Ebene. Gleichwohl können sich Routinen bzw. feststehende Strukturen ändern. Hierfür sind Interaktionen zwischen Individuen maßgeblich. Dieser Aspekt ist in Bezug auf den Gegenstand der Arbeit insofern von Relevanz, wenn es um die Frage geht, ob Inklusion als Innovation verstanden und ob es in der Praxis implementiert wird.

5.1.3 Sensemaking und Macht

Der Sensemaking-Ansatz wurde von verschiedenen Autor_innen dafür kritisiert, den Machtaspekt in Organisationen zu vernachlässigen (z. B. Mills et al. 2010): „Sensemaking strikes some people as naive with regard to the red meat of power, politics, and critical theory“ (Weick et al. 2005, S. 418). So könne die Umsetzung einer politischen Vorgabe nicht allein dadurch erklärt werden, welche Prozesse auf individueller sowie subjektiv-generischer Ebene ablaufen. Stattdessen wären Top-Down-Vorgaben bei der Implementation von politischen Vorgaben maßgeblich.

Weick et al. (2005) widersprechen diesem Argument: So müsse man sich im Rahmen des Sensemaking-Ansatzes mit den folgenden Fragen auseinandersetzen: „How does power get expressed, increase, decrease, and influence others“ (ebd., S. 418). Beispielsweise kann Macht ausgeübt werden, indem der Informationsaustausch von formell höherstehenden Akteuren beeinflusst wird. Des Weiteren zeigen sich Machtverhältnisse auch, wenn Interessen und Positionierungen von formell höherstehenden Akteuren berücksichtigt werden, ohne dass diese explizit Macht ausüben müssen.

In Bezug auf Machtverhältnisse argumentiert Weick darüber hinaus, dass Machtausübung nicht als Einbahnstraße verstanden werden kann, wonach etwas von „unten“ so umgesetzt wird, wie es von „oben“ intendiert war. Stattdessen wird den Sensemakern eine aktive Rolle zugeschrieben, da sie eigene Deutungen und Überzeugungen einbringen können. Dies kann zur Folge haben, dass die Interessen von formell höherstehenden Akteuren ignoriert bzw. bewusst umgangen werden. Zusammenfassend lässt sich demnach festhalten, dass Sensemaking kein hierarchiefreier Prozess ist. Jedoch ist bei der Analyse von Sensemaking-Prozessen zu berücksich-

tigen, dass formelle Hierarchien in Bezug auf tatsächlich vorhandene Machtverhältnisse nicht immer aussagekräftig sind. Diese Annahme teilt der Ansatz mit dem Educational-Governance-Ansatz (nähere Ausführung → Kapitel 3), der im Rahmen dieser Arbeit in Hinblick auf das Bildungssystem als sensibilisierende Perspektive genutzt wird.

5.2 Implementation von politischen Vorgaben

Im Folgenden werden nun die einzelnen Elemente des von Spillane et al. (2002, S. 412) entwickelten „Distributed-Cognition“-Ansatzes näher dargestellt. Dieser stellt eine Konkretisierung des Ansatzes von Weick bezüglich der Implementation von politischen Vorgaben dar. Zuvor wird jedoch dargestellt, inwiefern sich der Sensemaking-Ansatz von anderen theoretischen Ansätzen aus der Implementationsforschung unterscheidet.

5.2.1 Abgrenzung zu anderen Ansätzen

Der Sensemaking-Ansatz wird in den letzten Jahren vermehrt in der Forschung zur Implementation von politischen Vorgaben angewandt. Dabei bezieht sich der Begriff „politische Vorgabe“ unter anderem auf konkrete Gesetzesvorgaben, politische Maßnahmen, Rechtsverordnungen und Ähnliches.¹ In Anlehnung an die von Weick (1995) beschriebenen Prozesse des Sensemakings („creation“, „interpretation“ und „enactment“) finden bei der Deutung von politischen Vorgaben folgende Prozesse statt:

„Conceptualizing the problem of implementation in this way focuses attention on how implementing agents construct the meaning of a policy message and their own behavior, and how this process leads or does not lead to a change in how they view their own practice, potentially leading to changes in both understanding and behavior“ (Spillane et al. 2002, S. 392).

Angewandt auf den Gegenstand der Arbeit bedeutet dies, dass sich die Instrumenteentwickler_innen mit den folgenden Fragestellungen beschäftigen:

Creation: Welche Indikatoren bzw. „cues“ von Inklusion werden aufgegriffen, um Inklusion als politische Vorgabe greifbar und interpretierbar zu machen?

Interpretation: Welche möglichen Deutungen von Inklusion gibt es? Welche möglichen Formen der Messung von Inklusion gibt es?

Enactment: Was wird unter Inklusion verstanden? Was sind Instrumente zur Messung von Inklusion?

Verortung in der Implementationsforschung : Der Sensemaking-Ansatz ermöglicht durch seine sozialkonstruktivistische Perspektive aus mehreren Gründen eine Erweiterung zu anderen theoretischen Ansätzen aus der Implementationsforschung. So gehen viele Ansätze davon aus, dass lokale Akteure bei der Implementation von politischen Vorgaben sich dessen bewusst sind, welche Intentionen hinter einer politischen Vorgabe stehen und wie sie diese auf ihrer Handlungsebene umsetzen müssen (vgl. Spillane 2004, S. 6). Demnach sind nur drei Szenarien

¹ Im Rahmen dieser Arbeit geht es primär nicht um die Implementation von Art. 24 BRK, sondern darum, wie Inklusion als politische Vorgabe – wie sie in verschiedenen Verordnungen der UN Eingang gefunden hat – von Seiten der Inspektionen gedeutet wird (nähere Ausführung → Kapitel 2).

bei der Deutung von politischen Vorgaben möglich: **Erstens**, lokale Akteure setzen die politische Vorgabe so um, wie sie intendiert ist. **Zweitens**, sie ignorieren die politische Vorgabe bewusst und werden nicht aktiv oder **drittens**, sie deuten diese nach ihren eigenen Interessen um und ignorieren die ursprüngliche Intention der politischen Vorgabe. Dagegen berücksichtigt der Sensemaking-Ansatz auch als Option, dass die Deutung der politischen Vorgabe von verschiedenen Rahmenbedingungen geprägt wird und dies zur Folge haben kann, dass diese von den lokalen Akteuren – unbewusst bzw. aus Unwissenheit – nicht so gedeutet wird, wie sie intendiert war. Dies kann zugleich implizieren, dass eine politische Vorgabe von Seiten der Instrumenteentwickler_innen konkretisiert oder neu gedeutet wird (vgl. Coburn 2001, S. 145) und die Instrumenteentwickler_innen der Schulinspektionen somit selbst als Policymaker fungieren.

Spillane et al. grenzen sich demnach von Ansätzen wie dem Rational-Choice-Ansatz ab, wonach Akteure sich nur von ihren eigenen Interessen leiten lassen (vgl. Spillane et al. 2002, S. 391). Des Weiteren kritisieren sie auch institutionelle Erklärungsansätze, die den Fokus auf strukturelle Bedingungen legen und dabei individuelle Einflussfaktoren vernachlässigen. Demnach sieht der Sensemaking-Ansatz die Ursachen für bestimmte Deutungen in strukturellen Bedingungen, aber darüber hinaus können auch individuelle Attribute sowie informelle Settings, z. B. Netzwerke, den Deutungsprozess prägen (vgl. ebd., S. 414). Ein Mehrwert des Ansatzes ist darüber hinaus, dass auch auf Erkenntnisse aus der Kognitionspsychologie zurückgegriffen wird, um Deutungsprozesse zu erklären. Dabei geht der Ansatz davon aus, dass die Implementation einer politischen Vorgabe nicht isoliert von anderen Anforderungen geschieht, die an die deutende Person bzw. Organisation formuliert werden. Beispielsweise ist die Entwicklung von Messinstrumenten zu Inklusion eine von vielen Aufgaben der Instrumenteentwickler_innen.

Im Vergleich zu anderen Ansätzen distanziert sich der Sensemaking-Ansatz von einfachen Ursache-Wirkungs-Mechanismen und es wird angenommen, dass Deutungsprozesse komplexer sind und daher die Interaktion zwischen verschiedenen Rahmenbedingungen in den Blick genommen werden muss. Des Weiteren gibt der Ansatz Hinweise darauf, welche Rahmenbedingungen bei Deutungsprozessen von Relevanz sein könnten, aber er lässt auch zu, dass in der empirischen Analyse bislang unbekannte Bedingungen herausgearbeitet werden. Dies hat zur Folge, dass der Sensemaking-Ansatz durch eine große Offenheit gegenüber möglichen Rahmenbedingungen auf individueller oder struktureller Ebene gekennzeichnet ist. Hierbei wird zugleich ein Kritikpunkt am Sensemaking-Ansatz sichtbar: Auf die Frage danach, welche Rahmenbedingungen von Relevanz sind, kann der Ansatz als solcher nicht falsifiziert werden (nähere Ausführung folgt). Demnach kann der Sensemaking-Ansatz als Sozialtheorie bezeichnet werden, welche durch „empirische Forschung nicht falsifizierbar“ (Burzan und Hitzler 2017, S. 5) ist.

Theoretische Ansätze in der deutschsprachigen Bildungsforschung : In der deutschsprachigen Bildungsforschung ist der Sensemaking-Ansatz eher unbekannt (z. B. Hartung-Beck und Diemer 2009; Zeitler et al. 2013; B. Koch 2011). So werden in den erziehungswissenschaftlichen und soziologischen Arbeiten im deutschsprachigen Raum bezogen auf die Implementation von politischen Vorgaben häufig der Rekontextualisierungsansatz von Fend (2006) oder der Neoinstitutionalismus-Ansatz (vgl. Scott 2008) angewandt.

Der **Rekontextualisierungsansatz** (z. B. Amrhein 2011; siehe auch Hummel 2016; Lambrecht 2013) stimmt mit dem Sensemaking-Ansatz darin überein, dass nicht alles „was vom Gemeinwesen auf bildungspolitischer Ebene gewollt ist, auf unverfälschte Weise bei Lehrern und Schülern

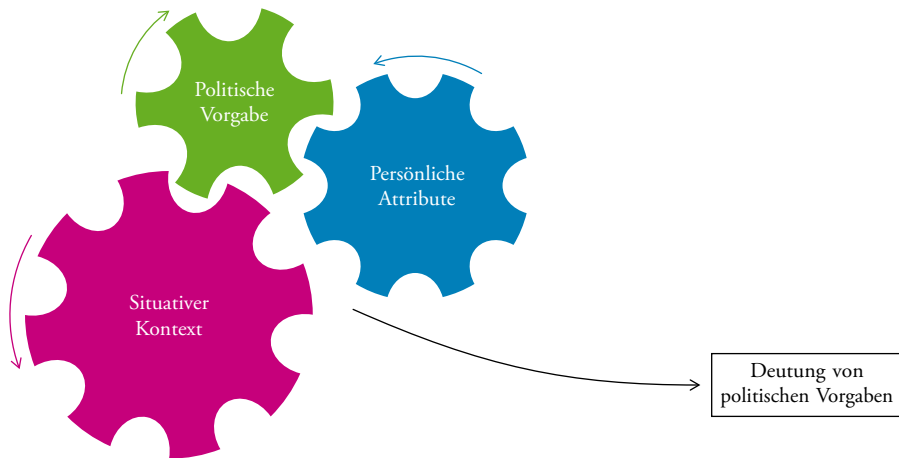


Abb. 5.2: Sensemaking von politischen Vorgaben

ankomme“ (Fend 2006, S. 175). Stattdessen werden die Vorgaben auf den einzelnen Ebenen des Bildungssystems, z. B. Bildungspolitik oder Unterrichtsebene, re-interpretiert und adaptiert (vgl. ebd., S. 174). „Diese Adaption ist neben den institutionellen Vorgaben von reflexiven Prozessen der Selbst- und Fremdwahrnehmung, von Kompetenzen der Aufgabenerfüllung und von situativen Konstellationen beeinflusst“ (ebd., S. 175). Beim konkreten Prozess der Rekontextualisierung fehlt es bei Fend jedoch zum Teil an genaueren Angaben: Es gibt keine Hinweise, welche Prozesse auf individueller Ebene ablaufen. Darüber hinaus mangelt es an einer genauen Beschreibung des Settings, in dem die Rekontextualisierung stattfindet. Diese Aspekte werden bei dem Sensemaking-Ansatz expliziter herausgearbeitet und ermöglichen dadurch eine differenzierte Auseinandersetzung bei der empirischen Analyse von Daten.

Im Gegensatz zum **Neoinstitutionalismus**, der in der Inklusionsforschung u. a. von Blanck et al. (2013) angewandt wird, ermöglicht es der Sensemaking-Ansatz – wie bereits oben dargestellt – auch die individuelle Ebene, d. h. persönliche Wertvorstellungen und Erfahrungen, in den Blick zu nehmen. Demnach kann es zwischen Individuen variieren, welche Aspekte der Umwelt für relevant angesehen werden. Innerhalb der Organisationsforschung kennzeichnet sich der Sensemaking-Ansatz daher durch eine konstruktivistische Epistemologie (vgl. Wetzel 2001, S. 163).

5.2.2 „Distributed-Cognition“-Ansatz

Der Sensemaking-Ansatz wurde von Spillane et al. (2002) in Bezug auf die Implementation von politischen Vorgaben konkretisiert. Nach dem von ihnen entwickelten „Distributed-Cognition“-Ansatz können die Rahmenbedingungen, die das Sensemaking von politischen Vorgaben prägen, drei Elementen zugeordnet werden: **den persönlichen Attributen des Sensemakers, dem situativen Kontext sowie der Repräsentation der politischen Vorgabe**. Indem sie diese drei Elemente benennen, konkretisieren sie den von Weick eher allgemein gehaltenen Begriff des Referenzrahmens (→ Abb. 5.2). Im Folgenden werden diese näher beschrieben.

Element I: Persönliche Attribute des Sensemakers: Spillane et al. (2002) argumentieren, dass die individuellen Kenntnisse, Überzeugungen und Erfahrungen des Sensemakers Einfluss darauf haben, wie diese die politische Vorgabe implementieren (vgl. ebd., S. 388). Beispielsweise kann bei der Deutung die fachliche Expertise der jeweiligen Sensemaker relevant sein (vgl. DeMatthews 2012). Des Weiteren spielen auch die Wertevorstellungen und Emotionen des Sensemakers eine Rolle (vgl. Spillane et al. 2002, S. 401). So argumentieren sie in Bezug auf den Bildungsbereich, dass viele Vorstellungen über Unterricht, Schule usw. stark normativ besetzt sind (vgl. ebd., S. 401) und somit nicht getrennt von den Wertvorstellungen des Sensemakers betrachtet werden können.

Element II: situativer Kontext : Als weiteres Element ist der situative Kontext zu nennen. Dabei wird das Konzept des situativen Kontexts sehr weit gefasst:

„Situation or context is a multifaceted construct that includes everything from national and professional identities to the structures of the offices and organizations in which people work. Implementing agents encounter policy in a complex web of organizational structures, professional affiliations, social networks, and traditions“ (ebd., S. 404).

Als konkrete Faktoren in Bezug auf den situativen Kontext benennen sie beispielsweise:

- **Setting der Auseinandersetzung:** Inwiefern haben die Akteure die Möglichkeit, sich über die jeweilige Vorgabe auszutauschen und ihr Handeln abzustimmen (vgl. ebd., S. 408). Fehlende Austauschmöglichkeiten können dazu führen, dass innerhalb einer Organisation unterschiedliche Verständnisse darüber existieren, wie die Vorgabe zu implementieren ist.
- Darüber hinaus können die (**hierarchischen**) **Verhältnisse zwischen Mitarbeiter_innen** des Netzwerks eine Rolle spielen (vgl. Spillane 2004; siehe auch Coburn 2006, S. 344).
- Ein weiteres Element kann sich auch aus dem sogenannten **non-formalem Netzwerk** (vgl. Spillane et al. 2002, S. 409) ergeben. So können Akteure, die nicht formell in den Implementationsprozess einbezogen werden, informell trotzdem den Deutungsprozess prägen. Dies ist beispielsweise der Fall, wenn Kontakte zu bestimmten Lobbyverbänden bestehen (vgl. ebd., S. 409).
- Auch der **historische Kontext** kann unter anderem in Form von Pfadabhängigkeiten relevant sein. Beispielsweise existierten bei den Schulinspektionen bereits Messinstrumente, bevor Inklusion im öffentlichen Diskurs im Fall von Deutschland thematisiert wurde. Bei der Implementation von Inklusion könnten demnach die bereits vorhandenen Messinstrumente prägend wirken.
- Aber auch die **vorhandene Zeit**, um eine Vorgabe zu implementieren, sowie die **vorhandenen finanziellen Ressourcen** können – neben anderen Faktoren – den Implementationsprozess prägen (vgl. Spillane 2004, S. 104).
- Als weiteren Faktor im Rahmen des situativen Kontexts nennen Spillane et al. die **vorhandenen Wertvorstellungen innerhalb einer Organisation** (vgl. Spillane et al. 2002, S. 411).

Die fehlende Konkretisierung des situativen Kontexts hat den Vorteil, dass verschiedene Aspekte berücksichtigt werden können, zum Beispiel historische Pfadabhängigkeiten oder strukturelle Abhängigkeiten. Die fehlende Konkretisierung birgt jedoch auch die Gefahr, dass in der empirischen Erhebung bestimmte Aspekte verloren gehen, weil sie nicht Teil eines vorab gebildeten Suchrasters sind und somit in der Datenerhebung nicht sichtbar werden. Im

Rahmen dieser Arbeit wird dieser Aspekt insofern berücksichtigt, als auf den Educational-Governance-Ansatz (→ Kapitel 3) und auf bewertungssoziologische Konzepte (→ Kapitel 4) zurückgegriffen wird. Sie erfüllen hier die Funktion einer sensibilisierenden Perspektive, durch die Rahmenbedingungen, die möglicherweise den Deutungsprozess geprägt haben, in der empirischen Analyse mitgedacht werden können.

Darüber hinaus ist beim Sensemaking-Ansatz zu beachten, dass nur begrenzt abstrakte Wissensbestände, z. B. unterschiedliche Episteme, sichtbar werden (vgl. Sandberg und Tsoukas 2015). Sie werden als naturgegeben angesehen und daher in empirischen Erhebungen selten explizit gemacht.

Element III: Politische Vorgabe : Als weiteres Element, welches das Sensemaking von politischen Vorgaben prägt, ist die „politische Vorgabe“ zu nennen. Spillane et al. argumentieren, dass sich dieses Element nicht nur auf konkrete Vorgaben z. B. Gesetze, sondern auch auf deren Rezeption durch andere bezieht, wie „verbal and written media, including regulations, directives, legislation, workshops, and pamphlets of various sorts“ (Spillane et al. 2002, S. 414; siehe auch Spillane 2004, S. 174). Dabei kann es sein, dass diese Rezeptionen bereits Ergebnisse von anderen Sensemaking-Prozessen darstellen und die ursprüngliche Intention der politischen Vorgabe bereits verloren gegangen ist (vgl. ebd., S. 176). Spillane (2004) verweist in diesem Zusammenhang auf das Kinderspiel „Stille Post“²: Hierbei wird eine Nachricht flüsternd von Kind zu Kind weitergegeben. Im Laufe der Zeit verändert sich häufig der Inhalt der Nachricht bzw. Teilinformationen gehen verloren. Dies kann auch der Fall sein, wenn eine politische Vorgabe innerhalb des Bildungssystems weitergegeben bzw. kommuniziert wird.

Für die politische Vorgabe ist auch von Relevanz, welche Anforderungen an das System bzw. an die betroffenen Akteure formuliert werden. So kann es sein, dass eine politische Vorgabe nur kleine Veränderungen abverlangt. Eine politische Vorgabe kann jedoch auch einen Paradigmenwechsel implizieren, welcher zur Folge hätte, dass von den betroffenen Akteuren erwartet wird, dass sie sich von bestehenden Annahmen und Handlungsmustern verabschieden müssen. Demnach müsste der eigene Referenzrahmen kritisch hinterfragt und angepasst werden (vgl. Spillane et al. 2002, S. 415). Hierbei ist jedoch zu beachten, dass hinter einer politischen Vorgabe zwar die Intention stehen kann, große Veränderungen auszulösen, aber Sensemaker häufig dazu neigen, an bestehenden Wissensbeständen und Annahmen festhalten. Dies kann implizieren, dass die deutenden Personen beim Sensemaking dazu tendieren, politische Vorgaben so zu interpretieren, dass sie in ihren bestehenden Referenzrahmen passen bzw. nur geringe Veränderungen beinhalten (vgl. Spillane 2004, S. 78).

Bei empirischen Erhebungen ist nicht nur zu berücksichtigen, wie die politische Vorgabe vom Sensemaker rezipiert wird, sondern auch, wie sie ursprünglich intendiert war. Die Auseinandersetzung mit der ursprünglichen Intention einer politischen Vorgabe ist notwendig, um die jeweilige Implementation durch die lokalen Akteure einordnen zu können. Spillane betont, dass beide Perspektiven betrachtet werden müssen (vgl. ebd., S. 175): Der alleinige Fokus auf Sensemaker könnte implizieren, dass die Implementation einer politischen Vorgabe ein Erfolg sei. Beispielsweise, weil sich die Akteure sehr bemüht haben, allen Interessen gerecht zu werden. Erst durch die Frage, was durch die Policy-Ebene überhaupt intendiert wurde, lassen sich Aussagen darüber treffen, ob das, was intendiert war, auch auf anderen Ebenen des Bildungssystems umgesetzt wurde. Angewandt auf den Gegenstand der Arbeit heißt dies, dass auch untersucht werden muss, ob Inklusion als politische Vorgabe so gedeutet wurde,

2 Im US-amerikanischen Kontext heißt das Spiel „telephone game“.

wie es intendiert war. Dies ist in dieser Arbeit nur begrenzt leistbar, weil es nicht um die Implementation eines konkreten Gesetzes geht, sondern Inklusion wird als Konzept verstanden, welches in verschiedenen Konventionen verankert ist. So gibt es keinen Konsens darüber, was mit Inklusion als politische Vorgabe intendiert ist. Daher dient das Verständnis von Inklusion, welches im Inklusionskapitel (→ Abschnitt 2.2) herausgearbeitet wurde, als Vergleichsfolie.

5.2.3 Zwischenfazit

Durch den Sensemaking-Ansatz ist es möglich, Deutungen von politischen Vorgaben und deren Implementation in der Praxis zu analysieren und dabei herauszuarbeiten, welche Rahmenbedingungen den Deutungsprozess geprägt haben. Dabei steht die Interaktion zwischen dem Referenzrahmen und der politischen Vorgabe im Mittelpunkt der Analyse. In Hinblick auf den Referenzrahmen haben Spillane et al. verschiedene Elemente herausgearbeitet: persönliche Attribute des jeweiligen Sensemakers, den situativen Kontext sowie die politische Vorgabe. Diese Elemente stellen abstrakte Beschreibungen dar. Daran anknüpfend zielt die empirische Analyse darauf ab herauszuarbeiten, welche Aspekte in Bezug auf die einzelnen Elemente von Relevanz sind. Dies können beispielsweise im Fall der „persönlichen Attribute“ bestimmte Wertüberzeugungen sein oder im Fall des „situativen Kontexts“ konkrete Akteure oder bestehende Strukturen. Diese Aspekte werden im Folgenden als „Rahmenbedingungen“ bezeichnet. Von diesen Rahmenbedingungen können in Bezug auf die Deutung der politischen Vorgabe unterschiedliche Impulse ausgehen. Des Weiteren kann es sein, dass in zwei Bundesländern ähnliche Rahmenbedingungen von Relevanz sind, z. B. das Kultusministerium als Akteur, aber von den jeweiligen Ministerien unterschiedliche Impulse ausgehen. Die Rahmenbedingungen bzw. die sich daraus ergebenden Impulse ermöglichen es zu verstehen, wie es zu unterschiedlichen Deutungen kommt bzw. wieso es einzelfallübergreifend Überschneidungen gibt.

Bezüglich der stattfindenden Mechanismen ist zu berücksichtigen: Auf der Grundlage des Sensemaking-Ansatzes können Handlungen nicht durch Ursache-Wirkungs-Mechanismen erklärt werden, sondern die Sensemaker nehmen die jeweiligen Rahmenbedingungen als etwas wahr, dem sie sich nicht entziehen können (vgl. Spillane et al. 2002). Demnach widerspricht es den Annahmen des Sensemaking-Ansatzes, konkrete Mechanismen, wie Rahmenbedingung X und Y lösen Deutung Z aus, herauszuarbeiten. Weick et al. versuchen dies deutlich zu machen, indem sie zwischen „interpretation“ und „choice“ unterscheiden: Im Fall von Sensemaking hat die deutende Person nicht den Eindruck, dass eine Wahlmöglichkeit (= choice) vorhanden ist, sondern in dem Setting erscheint die jeweilige Interpretation für den Sensemaker kohärent (vgl. Weick et al. 2005, 409f).

Im Rahmen des Kapitels wurden bereits diverse Kritikpunkte des Ansatzes thematisiert, z. B. die fehlende Konkretisierung des situativen Kontexts, vage Beschreibungen der Beziehungen zwischen den Rahmenbedingungen sowie die Herausforderung bei der Berücksichtigung von temporalen Aspekten im empirischen Design. Weitere Kritikpunkte, die sich im Rahmen der empirischen Erhebung und Auswertung herauskristallisierten, werden im folgenden Kapitel sowie im Diskussionskapitel thematisiert.

6 Empirisches Design

Im Rahmen der Arbeit wird zum einen empirisch untersucht, wie Inklusion als Konzept von einzelnen Mitarbeiter_innen der Schulinspektion bzw. in den Instrumenten der Schulinspektionen verstanden wird (→ Kapitel 7). Zum anderen wird analysiert, wie die Instrumenteentwickler_innen der jeweiligen Schulinspektionen zu den Bewertungsinstrumenten gekommen sind (→ Kapitel 8). Für die Beantwortung der Fragestellungen wurde auf zwei Datenquellen zurückgegriffen: Erstens wurden die Bewertungsinstrumente der Schulinspektionen ausgewertet (nähere Ausführung → Abschnitt 6.1). Zweitens wurden Interviews mit den Instrumenteentwickler_innen der Schulinspektionen geführt, was im zweiten Abschnitt dieses Kapitels (→ 6.2) näher beschrieben wird. Bei der Beantwortung der Fragestellungen gibt es in Bezug auf die Datenquellen Überschneidungen: So wird sowohl mit den Dokumenten als auch mit Hilfe der Interviews die Frage beantwortet, wie Inklusion gedeutet wird (→ Tab. 6.1).

Die Verwendung eines qualitativen Forschungsdesigns ergibt sich aus der Fragestellung und dem theoretischen Ansatz der Arbeit. So zielt die Arbeit auf die „verstehend-interpretative Rekonstruktion sozialer Phänomene“ (Döring und Bortz 2016, S. 63) ab und insbesondere bei den Interviews stehen die Deutungen und Perspektiven der Interviewpartner im Mittelpunkt des Interesses. Des Weiteren gibt es bislang kaum Informationen zu den Inhalten und dem Entstehungsprozess der Bewertungsinstrumente. Dies erfordert ein exploratives Vorgehen und legt daher ein qualitatives Forschungsdesign nahe.

6.1 Analyse der Bewertungsinstrumente

Durch die Analyse der Bewertungsinstrumente werden die folgenden Fragestellungen beantwortet:

- Inwiefern wird in den Instrumenten der Schulinspektionen ein Verständnis von guter Schule sichtbar, wonach im schulischen Alltag die Teilhabe und Nichtdiskriminierung aller (~ Inklusion) angestrebt wird?
- Welche Verständnisse von Inklusion spiegeln sich in denjenigen Instrumenten wider, die von Seiten der Schulinspektionen als Instrumente für die Messung von Inklusion gekennzeichnet sind?

Es folgen nun Informationen zur Stichprobe, Datenerhebung und zur Auswertung. Danach wird beschrieben, wie die Bewertungsinstrumente mit Hilfe der qualitativen Inhaltsanalyse nach Gläser und Laudel ausgewertet wurden.

6.1.1 Datenkollektion

Bei der Datenkollektion waren die folgenden zwei Kriterien von Relevanz: **Erstens**, es wurden nur die Instrumente derjenigen Schulinspektionen berücksichtigt, die zum Zeitpunkt der Datenerhebung, d. h. im Oktober 2015, Inspektionen durchgeführt haben bzw. deren Einführung absehbar war. So kündigte das Kultusministerium in Schleswig-Holstein zu diesem Zeitpunkt an, dass es im Februar 2016 eine Schulinspektion einführen würde. **Zweitens** die Bewertungsinstrumente waren öffentlich zugänglich und somit auch für Schulen einsehbar. Dieses Kriterium

Tab. 6.1: Datenmaterial

	Analyse der Bewertungsinstrumente , z. B. Unterrichtsbeobachtungsbögen	Analyse der Experteninterviews
Stichprobe	N = 12 Bundesländer bzw. Fälle	N = 18 Interviews in 9 Bundesländern
Betrachtungsebene	Auf Bundesländerebene	Individualebene (und z.T. Bundesländerebene)
Beantwortung der Fragestellungen	Deutung von Inklusion bzw. Instrumente zu Inklusion in den Bewertungsinstrumenten der Schulinspektionen	Deutung von Inklusion bzw. inklusiver Bewertungsinstrumente durch die Instrumenteentwickler_innen Rahmenbedingungen, die die Entwicklung der Instrumente geprägt haben
Verknüpfung zum Sensemaking-Ansatz	Ergebnis des Sensemakings im Rahmen von kollektiven Aushandlungsprozessen, in denen verschiedene Einzelpersonen involviert waren.	Ergebnis des Sensemakings auf individueller Ebene Beschreibung der Rahmenbedingungen, die den Prozess auf individueller Ebene sowie auf kollektiver Ebene geprägt haben
Auswertungsmethode	Qualitative Inhaltsanalyse nach Gläser und Laudel	Qualitative Inhaltsanalyse nach Gläser und Laudel

basierte auf dem Gedanken, dass Schulinspektionen gegenüber Schulen unter anderem eine normsetzende Funktion ausüben (nähere Ausführung → Kapitel 3). In Einzelfällen wurden die Schulinspektionen für konkrete Dokumente angefragt. Hierbei ist jedoch zu berücksichtigen, dass diese Instrumente auch für die Schulen auf Anfrage vorab einsehbar sind bzw. sie diese Dokumente erhalten, wenn eine Inspektion an ihrer Schule durchgeführt wird.

Zum Zeitpunkt der Datenerhebung gab es in 13 Ländern eine Schulinspektion (Bayern, Berlin, Brandenburg, Baden-Württemberg, Saarland, Bremen, Mecklenburg-Vorpommern, Hamburg, Hessen, Niedersachsen, Nordrhein-Westfalen, Sachsen-Anhalt, Schleswig-Holstein). Des Weiteren gab es noch die Bund-Länder-Inspektion¹ für deutsche Schulen im Ausland (blauer Kreis → Abb. 6.1). Kurz zuvor hatten drei Bundesländer die Schulinspektion abgeschafft bzw. diese pausiert. Dies war in Sachsen, Thüringen und Rheinland-Pfalz der Fall (→ Abb. 6.1). Daraus ergeben sich insgesamt 14 Fälle (13 Länder und die Bund-Länder-Inspektion), jedoch konnten nicht alle Fälle in die Analyse miteinbezogen werden. So waren im Fall von Bremen und Mecklenburg-Vorpommern die Bewertungsinstrumente nicht verfügbar. Nach Rücksprache ergab sich, dass diese Schulinspektionen zurzeit nicht systematisch mit standardisierten Bewertungsinstrumenten arbeiten. Dies hatte zur Folge, dass Bremen und Mecklenburg-Vorpommern

¹ Im Folgenden wird häufiger von „Bundesländern“ gesprochen. Damit ist auch die Bund-Länder-Inspektion gemeint, obwohl sie kein Bundesland repräsentiert.

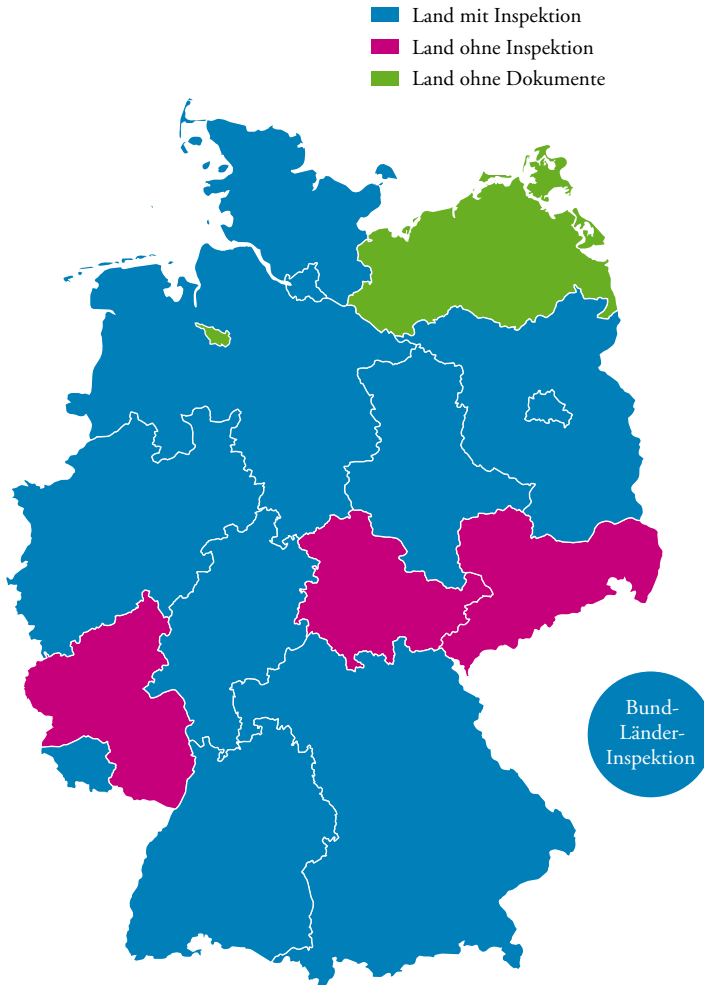


Abb. 6.1: Schulinspektionen in Deutschland

in der Stichprobe nicht berücksichtigt wurden. Für die Analyse der Instrumente bedeutet dies, dass insgesamt zwölf Fälle in die Analyse einbezogen wurden (→ Abb. 6.1).

6.1.2 Relevante Dokumente

Bei der Analyse der Instrumente wurde auf unterschiedliche Dokumententypen zurückgegriffen: Zu den relevanten Dokumenten gehören unter anderem Bewertungsinstrumente, wie die Unterrichtsbeobachtungsbögen und Fragebögen. Des Weiteren wurden im Rahmen der Analyse auch die Beschreibungen bzw. Merkmale von guter Schule, die unter anderem in den Handbüchern der Schulinspektionen aufgeführt werden, berücksichtigt.

Darüber hinaus greifen Schulinspektor_innen bei der Durchführung von Inspektionen auf interne Unterlagen zurück, z. B. ausführliche Erklärungen zu einzelnen Indikatoren. Wenn

diese nicht öffentlich zugänglich waren, wurden sie nicht berücksichtigt. Neben den internen Dokumenten werden die Schulqualitätsrahmen im Rahmen der Arbeit nicht als Bewertungsinstrumente verstanden und sind daher kein Teil der empirischen Analyse. Dies lässt sich dadurch erklären, dass die Schulqualitätsrahmen häufig sehr umfassende Beschreibungen von guter Schule enthalten. Diese sind für die konkrete Durchführung von Inspektionen kaum relevant. In den Instrumenten tauchen häufig nur Teilaspekte des Schulqualitätsrahmens auf auf bzw. bei der Operationalisierung bedarf es der Konkretisierung und Interpretation durch die Instrumenteentwickler_innen. Demnach werden die Schulqualitätsrahmen in Bezug auf die Entwicklung der Bewertungsinstrumente thematisiert (→ Kapitel 8), aber dieser werden im Rahmen der Arbeit nicht als Bewertungsinstrumente verstanden und sind somit nicht Teil der Analyse. Darüber hinaus konnten auch die so genannten Datenblätter zur Schulhausbegehung bei der Analyse nicht miteinbezogen werden, da diese nur in zwei Bundesländern zugänglich waren und sich daher nur begrenzt bundesländervergleichende Aussagen treffen lassen.

Im Rahmen der Analyse wird immer wieder explizit auf den jeweiligen Dokumententyp verwiesen. So weisen die einzelnen Dokumententypen unterschiedliche Logiken auf, was unter anderem damit zusammenhängt, dass die Dokumententypen dazu genutzt werden die Einschätzungen von unterschiedlichen Personen, z. B. von Eltern durch den Elternfragebogen oder von den Inspektor_innen durch den Unterrichtsbeobachtungsbogen, zu erfassen. Des Weiteren gibt es innerhalb des gleichen Dokumententyps nochmal Unterschiede zwischen den einzelnen Bundesländern, z. B. bei der Antwortskala. Auch dies wird im Rahmen der Analyse thematisiert, wenn es in Bezug auf die Fragestellung der Arbeit von Relevanz ist.

Da die Dokumente frei zugänglich sind, ist es möglich, die Bundesländer namentlich aufzuführen und miteinander zu vergleichen. Demnach findet sich hinter jedem Zitat die Nennung des Dokumententyps, z. B. „Unterrichtsbeobachtungsbogen (UBB)“ und die Abkürzung des jeweiligen Bundeslandes, z. B. Sachsen-Anhalt (ST). Zusätzlich ist zu berücksichtigen, dass die Bundesländer häufig unterschiedliche Begrifflichkeiten verwenden, um die jeweiligen Dokumententypen zu benennen. Zur Vereinheitlichung wird im Folgenden nur ein Begriff verwendet. Es folgt nun eine ausführlichere Beschreibung der einzelnen Dokumententypen.

„Merkmale von guter Schule“ im Handbuch (ÜT): Viele Schulinspektionen beschreiben im Rahmen eines Handbuchs oder auf der Website das jeweilige Verfahren der Inspektion, z. B. Ablauf oder Anzahl der Inspektor_innen. Im Handbuch befindet sich häufig auch eine Übersicht über die einzelnen Qualitätsbereiche von Schule und die dazugehörigen Indikatoren für gute Schule. Die Textbausteine, die beschreiben, was die jeweilige Schulinspektion unter einer guten Schule versteht (im Folgenden „Merkmale von guter Schule“), sind Teil der empirischen Analyse. Diese sind von den Messinstrumenten insofern abzugrenzen, als dass sie nicht zur konkreten Messung an Schulen genutzt werden, sondern sie können als Grundlegendokumente verstanden werden. Der Bezug zwischen diesen Merkmalen und den Instrumenten wird in einigen Bundesländern explizit hergestellt: Beispielsweise ist bei dem „Langprofil mit allen Profilmerkmalen und Kriterien“ des Landes Brandenburg vermerkt mit welchem Messinstrument das jeweilige Merkmal von guter Schule gemessen wird. So wird das Merkmal „Die Schülerinnen und Schüler zeigen sich zufrieden mit der Schule“ durch persönliche Interviews sowie durch den Schülerfragebogen abgefragt. In den Handbüchern mancher Bundesländer taucht eine Vielzahl von Beispielen auf, z. B. werden im Fall von Niedersachsen bei dem Qualitätsmerkmal „Kooperationen mit externen Partnern bestehen“ insgesamt acht Angebotsformate als Beispiele aufgeführt, wie „Gender Mainstreaming“; „Präventionsmaßnahmen“; Politische Bildung“ usw.

Diese Beispiele tauchen in den konkreten Instrumenten wie im Unterrichtsbeobachtungsbogen nicht mehr explizit auf. Hier könnte auf der einen Seite argumentiert werden, dass diese Aspekte bei der Messung „mitgedacht“ werden. Auf der anderen Seite besteht jedoch die Gefahr, dass diese Aspekte im Laufe der Zeit in Vergessenheit geraten. Dies ist bei der Interpretation der instrumentenübergreifenden Merkmale zu beachten und wurde in der Analyse durch die Kennzeichnung „Beispiele“ berücksichtigt (nähere Ausführung → Kapitel 7).

Darüber hinaus ist zu erwähnen, dass sich der Dokumententyp „Merkmale gute Schule“ im Format zwischen den Bundesländern unterscheidet: In den meisten Fällen ist es eine Aufzählung von Merkmalen von guter Schule und Indikatoren und/oder Beispielen, die nähere Hinweise zu den Merkmalen geben. Falls solch eine Aufzählung nicht auffindbar war, wurde auf deskriptive Beschreibungen des Verständnisses von Schulqualität zurückgegriffen, z. B. im Fall von Schleswig-Holstein. Insgesamt wurden die Merkmale von guter Schule in elf Fällen analysiert: Baden-Württemberg (BW), Bayern (BY), Berlin (BE), Brandenburg (BB), Hamburg (HH), Hessen (HE), Niedersachsen (NI), Nordrhein-Westfalen (NRW), Sachsen-Anhalt (ST), Schleswig-Holstein (SH) und Bund-Länder-Inspektion (BLI). Generell werden Zitate, die aus diesem Dokumententyp entstammen, in der Analyse mit dem Kürzel „ÜI“ („Übergreifende Indikatoren“) gekennzeichnet.

Unterrichtsbeobachtungsbogen (UBB): Ein wesentlicher Bestandteil einer Inspektion ist der Besuch des Unterrichts. Dafür besuchen die Inspektorinnen und Inspektoren mehrere Unterrichtsstunden und füllen hierbei den Unterrichtsbeobachtungsbogen aus. Auf den Bögen befinden sich verschiedene Items, die von den Inspektor_innen während der Unterrichtsbeobachtung ausgefüllt werden. Ein Item auf dem Berliner Unterrichtsbeobachtungsbogen ist beispielsweise folgender: „Die Raumgestaltung (auch die Sitzordnung) ist alters- und bedarfsgerecht“ (UBB_BE). Dabei unterscheiden sich die Bundesländer nicht nur in den Items, sondern auch in den Antwortskalen.

In Bezug auf die Antwortkategorien unterscheiden sich die Bundesländer in zwei Aspekten: Zum einen, ob sie eine dichotome Antwortskala anbieten (BW; BY; NI; NRW) oder zwischen vier Niveaustufen unterscheiden (BE; BB; HH; HE; SL ; ST) (→ Tab. 6.2). Zum anderen gibt es in einigen Bundesländern auch die Möglichkeit bei allen Indikatoren bzw. bei bestimmten Indikatoren anzugeben, dass dieser Aspekt nicht beobachtbar sei. Dahinter steht die Idee, dass eine Unterrichtsbeobachtung – außer in Bayern (Bayerisches Staatsministerium für Unterricht und Kultus 2010, S. 46) - 20 Minuten lang ist und somit möglicherweise nicht alle Aspekte von gutem Unterricht beobachtbar sind. Jedoch muss die Nicht-Beobachtbarkeit nicht implizieren, dass das Fehlen des Aspekts negativ zu bewerten ist. Im Fall von NRW ist dies beispielsweise in Bezug auf „Individuelle Lernwege“ der Fall: Falls in der Unterrichtsbeobachtung keine Form der Differenzierung sichtbar wird, müssen die Schulinspektor_innen auch nicht die dazugehörigen Indikatoren beantworten (UBB_NRW).

Für die Analyse des Unterrichtsbeobachtungsbogens wird auf 12 Fälle zurückgegriffen: Baden-Württemberg; Bayern; Berlin; Brandenburg; Hamburg; Hessen; Niedersachsen; Nordrhein-Westfalen; Saarland; Sachsen-Anhalt, Schleswig-Holstein und die Inspektion für Schulen im Ausland (BLI). Zitate aus dem Unterrichtsbeobachtungsbogen sind mit „UBB“ gekennzeichnet.

Fragebögen: Vor dem Besuch eines Schulinspektionsteams werden diverse Akteursgruppen gebeten, einen Fragebogen auszufüllen, der häufig im Rahmen einer Onlinebefragung realisiert

Tab. 6.2: Antwortkategorien

Bundesland	Antwortkategorien im UBB
Baden-Württemberg	„ja“ / „nein“; Bei manchen Indikatoren: „nicht einschätzbar“ (weil sie nicht in jedem Unterrichtsetting beobachtet werden können.
Bayern	In Bezug auf die Beobachtung: „(eher) ja“; „(eher) nein“; „keine Angaben möglich“
Berlin	„++“ / „+“ / „-“ / „--“ / Bei manchen Indikatoren: „#“ für nicht einschätzbar.
Brandenburg	„4 überwiegend stark“; „3 eher stark als schwach“; „2 eher schwach als stark“ und „1 überwiegend schwach“
Hamburg	„++ trifft voll zu“ / „+ trifft eher zu“ / „- trifft eher nicht zu“ / „0 nicht beobachtbar“
Hessen	„1 trifft nicht zu“ / „2 trifft eher nicht zu“ / „3 trifft weitgehend zu“ / „4 trifft in höchstem Maße zu“
Niedersachsen	„trifft zu“ / „trifft nicht zu“
Nordrhein-Westfalen	„trifft zu“ / „trifft nicht zu“
Saarland	„trifft zu“ / „trifft eher zu“ / „trifft eher nicht zu“ / „trifft nicht zu“ / hier nicht relevant
Sachsen-Anhalt	„1 trifft zu“ / „2 trifft eher zu“ / „3 trifft eher nicht zu“ / „4 trifft nicht zu“ / „0 konnte nicht beobachtet werden“

wird. Dieser zielt darauf ab möglichst repräsentative Daten zu erhalten. Dafür werden meistens Eltern, Lehrkräfte sowie Schüler_innen kontaktiert. Des Weiteren werden in manchen Bundesländern noch Erzieher_innen bzw. pädagogisches Personal befragt, z. B. in Berlin. Die Ergebnisse der Befragung dienen häufig als Grundlage, um im Rahmen von Einzel- oder Gruppeninterviews konkrete Nachfragen stellen zu können. Des Weiteren werden die Ergebnisse der Befragungen meistens auch in den Inspektionsberichten aufgeführt. Im Rahmen dieser Arbeit werden nur die Fragebögen für Eltern, Schüler_innen und Lehrkräfte analysiert, um eine Vergleichbarkeit zwischen den Bundesländern zu gewährleisten.²

Für die Analyse des Fragebogens für Lehrkräfte, Eltern und Schüler_innen wird auf folgende Fälle zurückgegriffen: Bayern, Berlin, Brandenburg, Hamburg, Hessen, Saarland, Sachsen-Anhalt sowie Schleswig-Holstein. Falls schulartspezifische Fragebögen vorhanden waren, wurde auf die Fragebögen für Oberschulen zurückgegriffen. Zitate aus den Fragebögen werden folgendermaßen gekennzeichnet: Beim Elternfragebogen „EFB“; beim Fragebogen für Lehrkräfte ist es die Abkürzung „LFB“ sowie beim Schülerfragebogen „SFB“.

Inspektionsbericht: Nach der Durchführung einer Inspektion erhalten die Schulen einen Inspektionsbericht. Die Inspektionsberichte geben Hinweise darauf, wie die erhobenen Daten aufbereitet und dargestellt werden. Dies hat wiederum Einfluss darauf, welche Interpretationen durch die Adressaten des Qualitätsurteils auf der Grundlage der Daten möglich sind. Daher

² Erhebungsformate wie die Einzelinterviews sind nicht standardisiert und somit im Rahmen dieser Arbeit nicht Teil der Analyse.

wurden exemplarisch einzelne Berichte hinzugezogen, wenn diese für das jeweilige Bundesland verfügbar waren. Jedoch waren in den meisten Bundesländern nur diejenigen Inspektionsberichte zugänglich, die von den Schulen selbst online veröffentlicht wurden. Lediglich im Fall von Hamburg, Berlin und Brandenburg wurde auf die Kurzversionen zurückgegriffen, die von den jeweiligen Ministerien veröffentlicht werden. Die Berichte werden im Folgenden mit „IB“ abgekürzt.

Zusätzliche Materialien in Bezug auf die explizite Verwendung des Inklusionsbegriffs:

In Bezug auf die Fragen, ob und wie der Inklusionsbegriff von Seiten der Schulinspektionen verwendet wird, wurden noch zusätzliche Materialien herangezogen, z. B. Schulqualitätsrahmen. Dies basiert auf dem Gedanken, dass bei diesem Teil der Analyse auch explizite Definitionen von Inklusion gesucht werden. Es ist jedoch eher unüblich, dass in konkreten Instrumenten Definitionen von Begrifflichkeiten aufgeführt werden.

6.1.3 Qualitative Inhaltsanalyse der Bewertungsinstrumente

Die Datenauswertung der Dokumente sowie der Experteninterviews orientiert sich am Verfahren der qualitativen Inhaltsanalyse nach Gläser und Laudel (2009). Im Vergleich zu anderen qualitativen Methoden weist das Verfahren aufgrund des regelgeleiteten Verfahrens Anknüpfungspunkte zu quantitativen Designs auf und ermöglicht es darüber hinaus viel Textmaterial zu bewältigen (vgl. Mayring und Fenzl 2014, S. 543). Wie bereits oben dargestellt, legt die Fragestellung der Arbeit ein qualitatives Design nahe. Die Entscheidung für die Inhaltsanalyse – im Vergleich zu anderen qualitativen Auswertungsverfahren – ergibt sich aus der theoretischen Herangehensweise sowie der Datenerhebung: Die qualitative Inhaltsanalyse kennzeichnet sich nach Mayring und Fenzl durch ihre „Kategoriegeleitetheit“ (ebd., S. 544). Demnach werden vor der Analyse der Texte bereits Kategorien („Codes“) entwickelt und die Analyse der Texte dient der Überprüfung, inwiefern die Kategorien im Material auftauchen, welche Ausprägungen der Kategorien vorhanden sind und welche Zusammenhänge sich herstellen lassen. Die Orientierung an bestehenden Kategorien bietet sich in Bezug auf die Fragestellungen der Arbeit an, da sich bei beiden empirischen Teilen, den Instrumenten sowie den Interviews, bereits vorab bestimmte Kategorien aus den theoretischen Vorüberlegungen ergeben. Ein weiterer Aspekt, der in Bezug auf die Auswahl der Auswertungsmethode von Relevanz ist, ist die Unterscheidung zwischen manifesten und latenten Sinnstrukturen: Als manifeste Sinnstrukturen werden Kommunikationsinhalte gesehen, die im Material explizit artikuliert werden. Dagegen erfordert das Herausarbeiten von latenten Sinnstrukturen eine Form der Interpretation, die über das tatsächlich Gesagte hinausgeht. Die qualitative Inhaltsanalyse erhebt zwar nach Mayring den Anspruch latente Sinnstrukturen herausarbeiten zu können (ebd., S. 543). Nach Lamnek ist es jedoch von der theoretischen Ausrichtung abhängig, ob mithilfe der qualitativen Inhaltsanalyse latentes oder manifestes Wissen herausgearbeitet werden kann (Lamnek 1995, S. 202). Auch wenn die qualitative Inhaltsanalyse den Zugriff auf latente Sinnstrukturen zulässt, ermöglichen es andere Auswertungsverfahren durch die intensive Auseinandersetzung mit einzelnen Textbausteinen stärker in die Tiefe zu gehen. Um Teilaspekte der Fragestellung dieser Arbeit näher zu betrachten, würde es sich bei zukünftigen Forschungsprojekten anbieten auf Auswertungsmethoden zurückzugreifen, die stärker latente Sinnstrukturen fokussieren, z. B. die objektive Hermeneutik. In Bezug auf den Gegenstand der Arbeit lässt sich jedoch festhalten, dass es kaum Forschung zur Entwicklung von Bewertungsinstrumenten gibt. Aufgrund eines

explorativen Designs spricht daher insbesondere für die qualitative Inhaltsanalyse, dass das gesamte Material berücksichtigt werden kann und somit auch Vergleiche zwischen einzelnen Bundesländern möglich sind.

Unter dem Begriff „qualitative Inhaltsanalyse“ gibt es eine Vielzahl von unterschiedlichen Verfahren. Im Folgenden wird dargestellt, warum sich das dargestellte Vorgehen an dem Verfahren nach Gläser und Laudel (2009) orientiert. Gläser und Laudel knüpfen bei ihrem Verfahren an dem Ansatz von Mayring (2010) an. Sie grenzen sich jedoch insofern davon ab, als dass sie während der gesamten Analyse und nicht nur nach der Betrachtung eines Teils des Materials die Möglichkeit zulassen Kategorien zu ergänzen (Gläser und Laudel 2009, S. 193). Nach Gläser und Laudel dürfen jedoch während der Analyse keine Kategorien aus dem ex-ante feststehenden Kategoriensystem entfernt werden (ebd., S. 193). Durch dieses Vorgehen ist daher ein induktiv-deduktives Vorgehen möglich. Dieses ermöglicht es während der gesamten Analyse Kategorien zu ergänzen und retrospektiv kritisch zu diskutieren, welche Kategorien in der Analyse keine Rolle gespielt haben. Aufgrund dieser Offenheit und der starken Orientierung an den theoretischen Vorüberlegungen orientiert sich die Arbeit im Folgenden an dem Verfahren der qualitativen Inhaltsanalyse nach Gläser und Laudel (2009), welches auf den folgenden Verfahrensschritten basiert:

- Theoretische Vorüberlegungen;
- Vorbereitungen der Extraktion;
- Extraktion;
- Aufbereitung und
- Auswertung (vgl. ebd., S. 203).

Im Folgenden wird die konkrete Anwendung der Verfahrensschritte zunächst für die Analyse der Instrumente und dann in Abschnitt 6.2 in Bezug auf die Interviews näher beschrieben. In diesem Zusammenhang wird auch erläutert, inwiefern das hier dargestellte Vorgehen von dem Verfahren nach Gläser und Laudel (2009) in Bezug auf Teilaspekte abweicht.

Vorbereitung der Extraktion bei der Analyse der Bewertungsinstrumente: Für die Extraktion ist die Entwicklung eines Kategoriensystems notwendig. Dieses stellt das „Herzstück“ (Schreier 2014) der Analyse dar und basiert auf den theoretischen Vorüberlegungen, die im theoretischen Teil dieser Arbeit ausführlich dargestellt wurden. Nach Gläser und Laudel dient die Entwicklung des Kategoriensystems dazu, „noch einmal das gesamte Wissen zu mobilisieren, um die Extraktion so gut wie möglich vorzubereiten“ (Gläser und Laudel 2009, S. 202).

In Bezug auf die Analyse der Instrumente geht es im Rahmen dieser Arbeit nicht darum eine Theorie zu überprüfen bzw. eine Theorie zu entwickeln. Somit weicht das Verfahren von der primären Zielsetzung nach Gläser und Laudel (2009) ab. Der Fokus liegt in diesem Auswertungsteil stattdessen auf der Frage, wie Inklusion in den Bewertungsinstrumenten verhandelt wird. Die Auseinandersetzung damit, wie Inklusion gedeutet wird, dient im weiteren Verlauf der Arbeit als Grundlage, um die Aussagen aus den Interviews zur Entwicklung der Instrumente einordnen zu können. Bei der Entwicklung des Kategoriensystems wird zwischen zwei Teilen unterschieden, da durch die Analyse der Instrumente zwei Fragestellungen beantwortet werden sollen.

1. **Berücksichtigung von Inklusion als normative Forderung:** Im ersten Schritt geht es um die Frage, ob bestimmte Konzepte, die auf einem menschenrechtlichen Inklusionsverständnis

Tab. 6.3: Kategoriensystem

Relevante Konzepte aus der theoretischen Ausarbeitung	Kategorien für die Analyse
Ziel von Inklusion: Individuell bestmögliche Qualifizierung	<ul style="list-style-type: none"> • Formen der Differenzierung • Lernstandserhebungen im Querschnitt
Ziel von Inklusion: Politische Partizipation	<ul style="list-style-type: none"> • Beteiligung aller Schüler_innen • Situationsspezifische Repräsentation von Minderheiten
Ziel von Inklusion: Erfahren von intersubjektiver Anerkennung	<ul style="list-style-type: none"> • Beziehungen zwischen schulischen Akteuren • Anerkennung von Individualität • Fehlen von Anerkennung, z. B. Exklusion, Diskriminierung
Berücksichtigung von sozial konstruierten Differenzlinien	<ul style="list-style-type: none"> • Relevante Differenzlinien • Operationalisierung von Differenzlinien

basieren, in den Instrumenten sichtbar werden und wie diese operationalisiert werden. So lässt diese Analyse Aussagen darüber zu, inwiefern Inklusion als politische Vorgabe – unabhängig von der expliziten Nennung des Begriffs – in den Instrumenten berücksichtigt wurde. Als Dimensionen dienen hierfür die Konzepte, die im theoretischen Teil formuliert wurden → Tab. 6.3).

2. **Verwendung des Inklusionsbegriffs in den Instrumenten:** Nun geht es um die Frage, welche Verständnisse von Inklusion sich in denjenigen Instrumenten widerspiegeln, die von Seiten der Schulinspektionen als „inklusionsorientiert“ gekennzeichnet wurden? Dabei wurden einzelne Definitionen des Inklusionsbegriffs extrahiert. Des Weiteren wurden induktiv Kategorien entwickelt, um zwischen einzelnen Instrumententypen unterscheiden zu können:

- **Integration der Instrumente in das bestehende Instrumentenset**, z. B. Add-On oder Querschnitt
- **Inhalte der Instrumente zu Inklusion**
- **Verständnis von Inklusion bei der expliziten Verwendung des Inklusionsbegriffs**

Extraktion und Aufbereitung bei der Analyse der Messinstrumente: Das Ziel des nächsten Verfahrensschritts, der Extraktion, ist es, aus den Dokumenten die Informationen zu entnehmen, die in Bezug auf die Fragestellung von Relevanz sind (vgl. ebd., S. 194). Bei der Entscheidung darüber, ob eine Textstelle codiert wird, dient als primäre Grundlage das Kategoriensystem. Darüber hinaus ist es jedoch stets möglich weitere Kategorien induktiv zu ergänzen. Zu Beginn wurde die Extraktion bei den Instrumenten mithilfe einer Software für qualitative Textanalyse durchgeführt. Hierbei zeigte sich, dass bei der Analyse mehrere Aspekte zeitgleich berücksichtigt werden mussten: **Erstens**, sollten Unterschiede zwischen den Bundesländern sichtbar werden. **Zweitens** sollten – unabhängig vom Bundesland – unterschiedliche Formen der Operationalisierung bei den einzelnen Kategorien herausgearbeitet werden. **Drittens** sollten auch unterschiedliche Formen der Operationalisierung in Bezug auf die einzelnen

ÜI: Erfahren von intersubjektiver Anerkennung

Bundesland	Wertschätzende / respektvolle Behandlung der Lehrkraft durch SUS	Wertschätzende / respektvolle Behandlung der SUS durch LK/SL	Respektvoller Umgang zwischen SUS	Respektvoller Umgang zwischen allen Akteuren	Fehlen von Gewalt vs. Diskriminierung bzw. Umgang damit	Anerkennung des Individuallität der Lehrkraft
Baden-Württemberg		ÜI_BW: „Die Lehrkräfte sowie Schülerinnen und Schüler gehen respektvoll und wertschätzend miteinander um.“ ÜI_BW: „Die Schulleitung zeigt Respekt, Anerkennung und Fürsorge im Umgang mit...“	ÜI_BW: „Die Schülerinnen und Schüler nehmen im Unterricht Rücksicht aufeinander.“ Indikatorebene ÜI_BW: „Die Schülerinnen und Schüler motivieren,...“	ÜI_BW: „Die Schule sorgt für den wertschätzenden Umgang innerhalb der heterogenen Schulgemeinschaft und fördert ihn aktiv.“ Indikatorebene ÜI_BW: „Alle an der Schulgemeinschaft Beteiligten gehen angemessen und wertschätzend mit...“	ÜI_BW: „Maßnahmen zur Konflikt- und Problemlösung sowie präventive Maßnahmen zur Begegnung von Agressivität und Mobbing werden unter Einbindung von SuS durchgeführt.“ Indikatorebene	ÜI_BW: „Die Schülerinnen und Schüler fühlen sich als individuelle Persönlichkeiten wahrgenommen und gefördert.“ ÜI_BW: „Die Schülerinnen und Schüler sehen sich in ihrer...“

Abb. 6.2: Screenshot der Auswertung

Dokumententypen, z. B. Beobachtungsbögen oder Fragebögen, abgebildet werden. Um die Komplexität abbilden zu können, wurde im weiteren Verlauf auf ein Textverarbeitungsprogramm zurückgegriffen. Dort wurden auf 110 Seiten einzelne Tabellen (siehe Abb. 6.2) erstellt, die jeweils nur die Information über eine Dimension des betrachteten Konzepts von Inklusion, z. B. „Individuell bestmögliche Qualifizierung“ (siehe „B“ in Abb. 6.2) enthielten und wie diese in einem Instrumententyp, z. B. im Lehrerfragebogen (siehe „A“ in Abb. 6.2), auf Bundeslandebene operationalisiert wurde. Demnach stellten die Spalten einer Tabelle die unterschiedlichen Formen der Operationalisierung dar (siehe „C“ in Abb. 6.2), z. B. unterschiedliche Formen der Differenzierung in Bezug auf das Ziel „individuell bestmögliche Qualifizierung“. Jedes Bundesland (siehe „D“ in Abb. 6.2) verfügte über eine Zeile und in den jeweiligen Zellen befand sich das Originalzitat bzw. eine Paraphrasierung.

Bei der Extraktion galten folgende Regeln:

- Bei der Durchführung war die Orientierung an den oben beschriebenen Dimensionen entscheidend. Es wurde vorab nicht festgelegt, welche Ausprägungen (siehe „C“ in Abb. 6.2) in Bezug auf die einzelnen Kategorien zu erwarten sind.
- Bestimmte Aspekte, z. B. Indikatoren zum Ganztagesbetrieb sowie Kriterien für Förderschulen, wurden in der Analyse nicht berücksichtigt. Die Nicht-Berücksichtigung von Förderschulen hat inhaltliche Gründe, da diese nicht Teil eines inklusiven Schulsystems sind. Indikatoren zum Ganztagsbetrieb wurden aus methodischen Gründen nicht einbezogen, um eine Vergleichbarkeit zwischen den Bundesländern zu gewährleisten.
- In Bezug auf das Ziel „individuell bestmögliche Qualifizierung“ waren nur die Indikatoren zu Lernprozessen relevant, die irgendeine Form von Differenzierung enthielten. Beispielsweise wurde das Item „Schüler sind weder unterfordert noch überfordert“ (UBB_ST) nicht aufgenommen, weil nicht zwischen einzelnen Schüler_innen bzw. Schülergruppen differenziert wurde.
- In Bezug auf den Unterrichtsbeobachtungsbogen (UBB) wurde in der Extraktion nicht die Erfassung von beobachteten Sozialformen zu Beginn des Unterrichtsbeobachtungsbogens,

z. B. Stattfinden von Gruppenarbeit, berücksichtigt. So war nicht ersichtlich, was von Seiten der Inspektor_innen als „gut“ oder „schlecht“ bewertet wird.

Auswertung bei der Analyse der Messinstrumente: In Bezug auf den vierten Verfahrensschritt, der Auswertung, machen Gläser und Laudel (2009) keine genaueren Angaben, da das Vorgehen abhängig von der jeweiligen Fragestellung ist. Die Auswertung der Bewertungsinstrumente zielt **im ersten Schritt** der Analyse darauf ab, zu untersuchen, ob die in Abschnitt 2 dargestellten Konzepte von Inklusion in den bundesländerspezifischen Instrumenten der Schulinspektionen eine Rolle spielen und welche Formen der Operationalisierung gewählt wurden. Dabei soll das Spektrum an unterschiedlichen Formen der Operationalisierung abgebildet werden. In Bezug auf Teilaspekte lassen sich zusätzlich Aussagen darüber treffen, inwiefern sich die Bundesländer voneinander unterscheiden. Demnach wird in Abschnitt 7.1 dargestellt, ob und wie die jeweiligen Konzepte in den einzelnen Bundesländern operationalisiert wurden. **Im zweiten Schritt** geht es darum zu beantworten, was unter Inklusion in den bundesländerspezifischen Instrumenten der Schulinspektionen verstanden wird und inwiefern es hier Abweichungen zu den in Abschnitt 2.2 beschriebenen Konzepten zu Inklusion gibt. Hierfür wird in Abschnitt 7.2 zunächst dargestellt, welche Verständnisse von Inklusion sich in den Instrumenten widerspiegeln und im nächsten Schritt wird diskutiert, inwiefern diese Gemeinsamkeiten oder Unterschiede zu dem in Abschnitt 2.2 beschriebenen Konzepten aufweisen. Die Auswertung wurde von der Autorin der Arbeit alleine durchgeführt. Jedoch wurden die Ergebnisse der Analyse durch eine studentische Hilfskraft überprüft und verifiziert.³ Aufgrund der Rückmeldungen wurden einzelne Punkte noch ergänzt bzw. angepasst.

6.2 Analyse der Experteninterviews

Zusätzlich zu der Analyse von Dokumenten wurden Experteninterviews geführt. Diese dienen zur Beantwortung der folgenden Fragestellungen:

- Was verstehen die Instrumenteentwickler_innen auf individueller Ebene unter Inklusion?
- Welche Rahmenbedingungen wirkten prägend?
- Inwiefern lassen sich Inklusionsdeutungen auf die Verschränkung von unterschiedlichen Rahmenbedingungen zurückführen?

In neun Bundesländern wurden teilstandardisierte Interviews mit Instrumenteentwickler_innen der Schulinspektionen durchgeführt. Die Entscheidung fiel auf **Leitfadeninterviews**, weil diese es ermöglichen, theoriegeleitet relevante Aspekte des Themas abzufragen (ebd.; Bernhard 2014, S. 100) und zugleich Offenheit für die Deutungen und Perspektiven der Interviewten zuzulassen. Jedoch waren die Fragereihenfolge sowie die Antwortmöglichkeiten nicht-standardisiert, so dass es möglich war flexibel auf die Antworten der Interviewpartner zu reagieren. Dadurch wurde die Annäherung an einen natürlichen Gesprächsverlauf angestrebt (vgl. Gläser und Laudel 2009, S. 42). Des Weiteren war es möglich, ad-hoc Nachfragen zu formulieren (vgl. ebd., S. 42). Der Fokus auf **Expert_innen** begründet sich nicht nur darin, dass sie eine „Quelle von Spezialwissen über die zu erforschenden sozialen Sachverhalte“ (ebd.) darstellen, „sondern auch (...) darin, dass dieses Wissen in besonderer Weise praxiswirksam und damit orientierungs-

³ In diesem Zusammenhang herzlichen Dank an Vera Klar-Winter für ihre Unterstützung bei der empirischen Analyse und ihre konstruktiven Rückmeldungen.

und handlungsleitend für andere Akteure wird“ (Bogner et al. 2014, S. 14). Dabei ist es unter anderem abhängig vom jeweiligen Forschungsinteresse, wer als Experte angesehen wird (ebd., S. 14). Bei der Auswahl der Gesprächspartner waren die Personen, die maßgeblich an der Entwicklung der Bewertungsinstrumente beteiligt waren, die Expert_innen. So verfügen sie über Wissen zu internen Prozessen und Strategien bei den jeweiligen Schulinspektionen. Dies ist insofern von Relevanz, als dass die Entwicklung von Bewertungsinstrumenten zumindest für die Öffentlichkeit nicht dokumentiert wurde und es dazu bislang kaum Forschung gibt. Demnach lässt sich durch die Interviews mit den Instrumenteentwickler_innen herausarbeiten, welche Rahmenbedingungen die Entwicklung geprägt haben.

6.2.1 Entwicklung des Leitfadens und theoretische Orientierung

Im Folgenden werden nun die Inhalte des Leitfadens dargestellt: Zu Beginn des Gesprächs wurden bestimmte Begrifflichkeiten geklärt, um Missverständnisse zu vermeiden, z. B. die Begriffe „Bewertungskriterien“ oder „Bewertungsinstrumente“. Des Weiteren wurden abhängig vom Bundesland bestimmte Begrifflichkeiten während der Interviews angepasst, z. B. wurden die Namen der jeweiligen Schulinspektionen und deren relevante Dokumente verwendet. Dies sollte dazu dienen dem Gesprächspartner zu signalisieren, dass die Interviewerin mit der jeweiligen Organisation vertraut ist. Nach der Einführung startete der Leitfaden mit zwei erzählgenerierenden Fragen: *„Können Sie sich kurz vorstellen und Ihre Rolle bei der [Name der Organisation in dem jeweiligen Bundesland] beschreiben?“* und *„Wie Sie wissen, interessiert mich der Entstehungsprozess der Bewertungskriterien. Können Sie mir mehr über den Entstehungsprozess erzählen?“*. Die erzählgenerierenden Fragen zu Beginn des Interviews wurden bewusst gewählt, um eine vertraute Gesprächsatmosphäre zu schaffen und der interviewten Person die Möglichkeit zu geben, eigene Schwerpunkte zu setzen (→ Abschnitt A im Anhang).

Es folgten drei Frageblöcke: **Im ersten Frageblock** ging um die Entwicklung der Bewertungsinstrumente im Allgemeinen. Da es kaum Literatur zu diesem Aspekt gab, wurde jenes Vorgehen gewählt, um einen Einblick in die üblichen Abläufe und Routinen bei der Entwicklung von Instrumenten zu erhalten. Hierbei orientierten sich die Fragen insofern am Sensemaking-Ansatz, als dass es konkrete Fragen zu den drei Elementen gab, die nach Spillane et al. (2002) den Deutungsprozess prägen: „persönliche Attribute des Sensemakers“, „situativer Kontext“ sowie „politische Vorgaben“. **Im zweiten Block** wurden Fragen gestellt, die an bewertungssoziologische Konzepte anknüpfen, z. B. „Welches Verständnis von guter Schule ist hier vorhanden?“. So geht es im Rahmen der Arbeit zwar primär um die Auseinandersetzung darüber, wie sich die Instrumenteentwickler_innen mit dem Konzept der Inklusion auseinandergesetzt haben. Da es jedoch um Instrumente geht, die zur Messung von Schulqualität genutzt werden, ist der Gegenstand der Arbeit nur schwer von der Frage zu trennen, was unter einer „guten“ Schule verstanden wird. Zusätzlich wurde auch erfasst, wie die Instrumenteentwickler_innen eine „schlechte“ Schule beschreiben würden (vgl. Lamont 2010). **Im dritten Frageblock** ging es um die Deutung von Inklusion und spezifische Rahmenbedingungen bei der Auseinandersetzung mit Inklusion. Durch indirekte Fragen wurde zusätzlich versucht zu erfassen, welche Relevanz dem Thema Inklusion beigemessen wird und welches Verständnis von Inklusion dominant ist: *„Könnte eine Schule, die nicht inklusiv ist, bei der Schulinspektion gut abschneiden?“*.

6.2.2 Kontaktaufnahme zu den Interviewpartnern

Bei der Kontaktaufnahme wurde den Schulinspektionen mitgeteilt, dass Personen gesucht werden, „die maßgeblich am Entstehungsprozess der Bewertungskriterien beteiligt waren“. Inhalte des Interviews seien die Entwicklung der Bewertungsinstrumente im Allgemeinen sowie Inklusion. Zusätzlich formulierte die Autorin dieser Arbeit gegenüber den Schulinspektionen die Bitte, dass sie gerne mit mind. zwei Personen sprechen würde. Dies führte dazu, dass die Autorin mit Personen sprach, die sich innerhalb der Schulinspektion für das Thema Inklusion zuständig fühlten, aber auch mit Personen, die an der Entwicklung der Instrumente beteiligt waren, aber nicht speziell für den Themenbereich Inklusion zuständig sind bzw. waren.

6.2.3 Durchführung der Interviews

Es wurden Einzelinterviews in 9 von 12 möglichen Bundesländern realisiert. Alle Interviews wurden persönlich an der jeweiligen Arbeitsstätte der interviewten Person durch die Autorin dieser Arbeit durchgeführt und nach Zustimmung der Interviewperson mit einem Aufnahmegerät aufgezeichnet. Dabei dauerten die Interviews ca. zwischen 30 und 90 Minuten.

In Bezug auf die Auswahl der Bundesländer ist festzuhalten, dass es keine Vorauswahl gab, sondern alle Fälle kontaktiert wurden. Bei vielen Bundesländern bestand bereits ein persönlicher Kontakt zum Feld.

Bei anderen Bundesländern gab es auch nach mehreren Anfragen keine Antwort auf eine Interviewanfrage beziehungsweise erst Monate später. Auch wenn kein Bundesland die Teilnahme an einem Interview explizit verneint hat, gab es einige wenige Bundesländer, die auf die Interviewanfrage abweichend reagierten. Im Juni 2016 kam es zur Entscheidung die fehlenden drei Fälle nicht mehr zu verfolgen und die Datenerhebung nach 18 Interviews abzuschließen. Der primäre Grund war die theoretische Sättigung, die sich bereits nach der Durchführung von ca. 13 bis 14 Interviews in sechs Fällen zeigte. Durch die Analyse der Bewertungsinstrumente ergab sich, dass zwei Bundesländer, die bislang noch nicht befragt worden waren, in Bezug auf einzelne Aspekte des Verfahrens bzw. der Bewertungskriterien stark abwichen. Daher wurden diese Fälle nochmals kontaktiert, da davon auszugehen war, dass sich jene Abweichungen möglicherweise auch darin widerspiegeln könnten, unter welchen Rahmenbedingungen die Instrumente entstanden sind. Des Weiteren meldete sich noch ein Bundesland verspätet auf die erste Kontaktanfrage, so dass insgesamt noch drei weitere Fälle befragt wurden. Abschließend wurden daher Gespräche in neun von zwölf möglichen Fällen durchgeführt. Demnach gibt es im ersten empirischen Teil Analysen von drei Fällen, die im zweiten empirischen Fall nicht im Rahmen von Interviews befragt worden sind. Eine Vollerhebung wäre aufschlussreich gewesen, wenn es die Möglichkeit gegeben hätte, die Analyse der Instrumente (→ Abschnitt 7.1 und 7.2) mit den Ergebnissen der Interviews (→ Abschnitt 7.3 und Kapitel 8) in den einzelnen Fällen zu verknüpfen. Dies war jedoch aus Datenschutzgründen nicht möglich. Da darüber hinaus bereits eine theoretische Sättigung eingetreten war, wurde von einer Vollerhebung abgesehen.

Insgesamt wurden 18 Interviews in neun Ländern geführt. Demnach wurde in jedem Bundesland durchschnittlich mit ca. zwei Personen persönlich gesprochen. Jedoch gab es Abweichungen, z. B. wenn es in dem jeweiligen Bundesland von Seiten der Schulinspektion nur möglich war ein Interview zu führen. So wurden in drei Fällen nur Interviews mit einer Person geführt, in vier Fällen mit zwei Personen, in einem Fall mit drei sowie in einem anderen Fall Interviews mit vier Personen durchgeführt. Die ungleiche Verteilung an Interviewpartnern ist insofern

nicht problematisch, als dass bei diesem Teil der Analyse die Individualebene und nicht die Bundeslandebene im Vordergrund steht.

Von den 18 Gesprächspartnern wurden zehn Personen von der Autorin als Frauen und acht als Männer gelesen. Die befragten Personen waren ca. zwischen 35 – 65 Jahre alt. Bis auf eine befragte Person hatte kein Interviewpartner sichtbare körperliche Beeinträchtigungen. Des Weiteren vermutet die Autorin, dass keine der Interviewten von anderen Personen als „Nicht-Deutsche“ markiert werden würde.⁴ Fast alle befragten Personen hatten nach eigenen Angaben Erfahrung als Lehrkraft gesammelt, bzw. waren neben ihrer Tätigkeit in der Schulinspektion noch zum Teil als Lehrkraft tätig.⁵ Fünf der 18 befragten Personen definierten sich selbst als „Expert_innen“ für Inklusion und fühlten sich innerhalb der Schulinspektion für das Thema zuständig. Vier Personen hatten die Leitung der jeweiligen Schulinspektion inne.

Nach der Durchführung der Interviews wurde das gesamte Audiomaterial transkribiert. In Bezug auf die Transkription dienten die Transkriptionsregeln (→ Abschnitt B im Anhang), die die Autorin im Rahmen eines anderen Forschungsprojekts (vgl. Piezunka et al. 2017) formuliert hatte, als Grundlage.⁶ Die 18 Interviews wurden von zwei studentischen Hilfskräften⁷ und von der Autorin der Arbeit transkribiert.

6.2.4 Reflexion der Interviews

Im Rahmen einer kritischen Reflexion geht es im Folgenden darum zu diskutieren, welche Aspekte den Verlauf des Interviews sowie die Auswahl der Informationen, die kommuniziert wurden, geprägt haben. Hierbei wird zum einen auf die Arbeiten von Becker (2006) sowie Alvesson (2003) zurückgegriffen. So wurden die Erkenntnisse beispielsweise in der Datenauswertung bei der Gewichtung von widersprüchlichen Aussagen berücksichtigt. Zum anderen dienen diese Informationen dazu, das vorhandene Material in Hinblick auf seine Aussagekraft einordnen zu können. Nachfolgend werden daher verschiedene Aspekte thematisiert, wie das Verhältnis zwischen Interviewerin und interviewter Person, Konstruktion von Identitäten, strategisches Verhalten durch die interviewte Person, Unsicherheit sowie geringe Thematisierung von persönlichen Erfahrungen.

Wahrnehmung der Interviewsituation: Nach Alvesson stellen Interviews „komplexe soziale Situationen“ (ebd., S. 18) dar. Hierbei spielt u. a. auch eine Rolle, wie die Interviewerin – in diesem Falle die Autorin dieser Arbeit – von den interviewten Personen wahrgenommen wird. Bei manchen Interviews hatte vorher bereits persönlicher Kontakt bestanden, z. B. inhaltlicher Austausch im Rahmen von Tagungen. Dies spiegelte sich in vielen (aber nicht in allen) Interviews darin wider, dass recht schnell eine sehr vertraute Gesprächsatmosphäre vorhanden war. Die vertraute Gesprächsatmosphäre und die Sympathie gegenüber den interviewten Personen stellte in diesbezüglich der Datenauswertung für die Autorin eine Herausforderung dar, da

4 Nach dem Sensemaking – Ansatz spielen auch persönliche Erfahrungen eine Rolle, wenn Konzepte wie Inklusion gedeutet werden. In diesem Zusammenhang könnte argumentiert werden, dass in Bezug auf einen weiten Inklusionsbegriff die oben aufgeführten Attribute des Sensemaker eine Rolle spielen. Jedoch ist es sehr schwer die oben genannten Merkmale zu erfassen. So ist zum einen zwischen Fremd- und Selbstidentifikation unterscheiden, d.h. ob die Person sich selbst als weiblich identifiziert oder ob der Person dieser Status zugeschrieben wird. Die dargestellte Beschreibung der Stichprobe stellt einen Versuch dar die Interviewpartner zu beschreiben und basiert auf Fremdzuschreibungen durch die Autorin dieser Arbeit.

5 Dies wurde nicht systematisch erfasst, sondern ergab sich aus den Angaben während des Interviews.

6 Auch hier wurden die Interviews nach Gläser/Laudel (2009) ausgewertet.

7 Ganz herzlichen Dank an Sophie Drebes und Tina Schaffus.

es ein langwieriger Prozess war, sich von den ersten oberflächlichen Interpretationen und Zuschreibungen zu lösen und eine Distanz zum Interviewmaterial aufzubauen.

Des Weiteren hatte die Autorin den Eindruck, dass ihr Geschlecht, Expertise und Alter im Gespräch mit einigen Interviewten eine Rolle spielten. So hatten manche Interviewpartner vor dem Interview bereits einen allgemeinen Vortrag der Autorin zur Umsetzung von Inklusion in Deutschland gehört. Dies hatte zur Folge, dass der Autorin von manchen Interviewpartnern ein „Expertenstatus“ in den Interviews zugeschrieben wurde, weshalb es während der Interviews zu interessierten Nachfragen kam, wie die Autorin dieser Arbeit bestimmte Aspekte einschätzen würde. Während des Interviews ist die Autorin diesen Fragen ausgewichen.

Jedoch gab es auch gegenteilige Erfahrungen: Als „junge“ Frau in der Qualifizierungsphase gab es einzelne Momente, wo die Autorin den Eindruck hatte, dass insbesondere einige männliche Interviewpartner eine väterliche Rolle einnahmen. Dies hatte häufig zur Implikation, dass im erklärenden Stil relativ viel und offen erzählt wurde. Es gab jedoch auch Situationen, in denen bestimmte problematische Themen bewusst umgangen wurden, weil davon ausgegangen wurde, dass die Problematik der Themen der Autorin dieser Arbeit nicht bekannt sei. Das Nachhaken der nicht angesprochenen Themen zum Ende des Interviews, führte dann – nach Eindruck der Autorin – zu Überraschungsmomenten. In diesem Zusammenhang ist es möglich, dass weitere Themen explizit nicht angesprochen wurden und dies aufgrund fehlender Kenntnisse durch die Autorin nicht bemerkt wurde.

Konstruktion von Identitäten: Nach Alvesson (2003) können Interviews als Instrument für die Interviewten dienen, um eigene Identitäten zu konstruieren bzw. zu „reparieren“. Dies spiegelt sich unter anderem darin wider, dass bei den Erzählungen unbewusst eine Rolle spielt, welche Identität sie von sich selbst erzeugen möchten. Dies kann unter anderem implizieren, dass Fehlverhalten nicht als solches gesehen und dementsprechend anders dargestellt wird. Inwiefern die Interviews zur Konstruktion oder zur Reparatur von Identitäten genutzt wurden, zeigte sich teilweise beim Vergleich der Interviews.

Die Mehrheit der Interviewten sprach zwar von Herausforderungen bei der Entwicklung der Bewertungsinstrumente, aber den finalen Bewertungsinstrumenten wurden kaum Mängel attestiert. So gab es nur wenige Fälle, in denen die Instrumente bzw. die Entwicklung kritisch reflektiert wurde: So gab es einen Fall, in dem der Status der Inspektion unsicher war (keine weiteren Informationen aus Anonymisierungsgründen) und einen anderen Fall, wo die Instrumente nicht mehr angewandt wurden. In Bezug auf die Datenauswertung ist der Vorgang der Identitätskonstruktion nur schwer zugänglich, da es kaum abschätzbar ist, welche Implikationen dies hinsichtlich des vorhandenen bzw. fehlenden Datenmaterials ergeben.

Politische Interessen bzw. strategisches Verhalten : Die Erzählungen der Interviewten sind auch davon geprägt, welche Botschaften sie der Interviewperson gerne mitgeben würde. Dabei kann das Kommunizieren von Botschaften sowohl unbewusst als auch intentional stattfinden (vgl. ebd.).

In einzelnen Interviewsituationen entstand der Eindruck, dass bewusst bestimmte Positionen vertreten bzw. Informationen weggelassen wurden. Zum Beispiel: Das Interview mit einer Leitungsperson war aus gesundheitlichen Gründen nach hinten verschoben worden. Die anderen Interviews, eine Präsentation zu den bundesländerspezifischen Ergebnissen durch die Autorin dieser Arbeit sowie eine anschließende Plenumsdiskussion hatten in dem jeweiligen Bundesland bereits eine Woche zuvor stattgefunden. Im Rahmen der Präsentation und der anschließenden

Diskussion wurde deutlich, dass die Mehrheit der Mitarbeiter_innen der Schulinspektion dem Thema Inklusion kritisch gegenübersteht. Beispielsweise wurde die Präsentation der Autorin mit den Worten angekündigt „Es wird heute ein Vortrag über ein Thema gehalten, das uns auch noch heimsuchen wird.“ In der Diskussion kam die Frage auf, warum Inklusion überhaupt notwendig sei und was so schlecht an den Förderschulen sei. Die hier beschriebene Interviewperson war zwar nicht anwesend, aber wies zu Beginn des Interviews darauf hin, dass sie das Protokoll gelesen habe und von den Mitarbeiter_innen informiert worden sei. In dem Interview fiel eine Vielzahl von Aussagen, die mit den Kenntnissen, die die Autorin im Rahmen der anderen Interviews, der Diskussion sowie durch die Analyse der Instrumente erworben hatte, nicht kohärent war. So wurde Inklusion zum Teil als ein Thema dargestellt, dass für die Mitarbeiter_innen der Schulinspektion relevant und unterstützenswert sei. Demnach kam es innerhalb des Interviews zu widersprüchlichen Aussagen dazu, welche Relevanz dem Thema beigemessen wird. Zusätzlich gab es noch von Seiten der Interviewperson den Hinweis, dass es vom Kultusministerium in den letzten Tagen das Signal gegeben hätte, dass das Thema Inklusion nun auf die Agenda müsste. Während des Interviews entstand der Eindruck, dass die Interviewperson aus strategischen Gründen eine bestimmte Position vertritt, wonach Inklusion eine große Bedeutung zugemessen wurde. Diese Position wurde jedoch im Laufe des Interviews nicht konsequent aufrecht gehalten. Für die Datenanalyse hatte diese Einschätzung zur Folge, dass einzelne Aussagen im Interview unterschiedlich gewichtet und eingeordnet wurden.

Unsicherheit bei den Interviewpersonen: Bei der Analyse des Interviews ist zu berücksichtigen, dass die Interviewsituation eine Ausnahmesituation darstellt, die häufig mit bestimmten Unsicherheiten einhergeht. Beispielsweise fragen sich Interviewte häufig, ob die Fragestellungen, die von den Forscher_innen kommuniziert wurden, die tatsächlichen Fragestellung sind (vgl. Alvesson 2003). In diesem Fall wurde die Fragestellung den Interviewten offen kommuniziert, aber dies kann den Interviewten nicht die Unsicherheit darüber nehmen, ob dies der Wahrheit entspricht. So zeigte sich in einem Fall, dass der Interviewpartner bereits im Vorhinein Nachfragen über das geplante Vorgehen bzw. während des Interviews mehrfach erkundigte, wie bestimmte Fragen gemeint seien. Bei der Beantwortung der Forschungsfragen nahm sich die Person sehr viel Zeit und wägte häufig ab, ob er/sie in der Lage sei, bestimmte Fragen zu beantworten bzw. verwies auf Kolleg_innen. Des Weiteren hatte dies zur Folge, dass der Erzählfluss nur schwer zustande kam. Dieses Vorgehen erklärte sich die Autorin zum einen mit Interesse an wissenschaftlichen Fragestellungen und dem Anspruch der interviewten Person als Berufsanfänger_in keine Fehler machen zu wollen. Zum anderen hatte die Autorin bei den Interviews den Eindruck, dass Unsicherheiten in Bezug auf die Zielsetzung des Interviews bei den anderen Interviewpartnern kaum vorhanden waren, bzw. im Laufe des Gesprächs abgebaut wurden.

Geringe Thematisierung von persönlichen Erfahrungen: Im Rahmen des Sensemaking-Ansatzes werden die persönlichen Überzeugungen und Erfahrungen des jeweiligen Sensemakers als ein Element genannt. So prägen diese die kognitive Auseinandersetzung mit Inklusion. Bei der Analyse der Interviews zeigte sich, dass ein Teilaspekt von persönlichen Überzeugungen, nämlich Aspekte auf der Gefühlsebene sowie Erfahrungen mit Inklusion im privaten Bereich, keine große Rolle gespielt haben. Es gibt verschiedene Erklärungen, die Hinweise darauf geben können, warum diese eine geringere Rolle gespielt haben, als dies vorab von der Autorin erwartet wurde: Eine Möglichkeit wäre, dass persönliche Erfahrungen bei der kognitiven

Auseinandersetzung mit Inklusion keine Rolle spielen. Dies lässt sich im Rahmen meiner Daten nicht falsifizieren. Dennoch ist diese Option bei einem normativ besetzten Konzept wie Inklusion eher unwahrscheinlich.

Eine weitere Möglichkeit ist, dass dieser Aspekt im Rahmen des Interviews kaum explizit thematisiert wurde: So ist bei der retrospektiven Betrachtung des Leitfadens sichtbar geworden, dass es nur wenige Fragen gab, die diesen Aspekt ansprechen. In diesem Zusammenhang sind jedoch zwei Aspekte zu nennen, die die Erklärungsmacht des Arguments zum Teil schwächen: So kennzeichnet sich der Leitfaden dadurch, dass recht allgemeine Fragen formuliert wurden bzw. viel Freiraum für Rahmenbedingungen gelassen wurde, die von Seiten der Interviewten eingebracht wurden. Der vorhandene Freiraum spiegelt sich beispielsweise in Bezug auf das Element „situativer Kontext“ in der Datenauswertung insofern wider, als dass eine Vielzahl von Akteuren und Rahmenbedingungen herausgearbeitet wurden, die bei der Entwicklung des Leitfadens noch nicht bekannt waren. Danach könnte argumentiert werden, dass durchaus die Möglichkeit bestanden hätte, Erfahrungen zu Inklusion aus dem privaten Bereich einzubringen, auch wenn diese nicht explizit abgefragt wurde. Ein weiterer Grund für die nicht allzu große Thematisierung könnte sein, dass die Interviewten in ihren professionellen Rollen angesprochen wurden. In diesem Zusammenhang wäre eine mögliche These, dass sie Erfahrungen aus dem privaten Bereich oder persönliche Überzeugungen von ihrer professionellen Rolle abgrenzen wollten und es daher kaum thematisiert wurde. Es stellt sich die Frage, wie diese Problematik bei zukünftigen Erhebungen umgangen werden könnte: Möglicherweise wären hierfür explizite Nachfragen nach Erfahrungen aus dem privaten Bereich notwendig. Abschließend lässt sich festhalten: Auch wenn Informationen zu Erfahrungen zu Inklusion aus dem privaten Bereich fehlen, gibt es eine Vielzahl von Rahmenbedingungen, die der Dimension „persönliche Attribute des Sensemakers“ zugeordnet werden können.

Positionierung und Expertise zu Inklusion von Seiten der Interviewten: Bei der Deutung von Inklusion ist ein Aspekt, welche Haltung gegenüber Inklusion eingenommen wurde. Hier stellt sich die Frage, inwiefern die Auswahl an Gesprächspartnern dadurch verzerrt war, dass die Autorin bei der Interviewanfrage an die Schulinspektionen bereits angegeben hatte, dass „Bewertungskriterien und Inklusion“ die Themen des Interviews sein werden. So liegt die Vermutung nahe, dass die Interviewpartner im Vergleich zu anderen Mitarbeiter_innen der Schulinspektion mit hoher Wahrscheinlichkeit im Durchschnitt positiver gegenüber Inklusion eingestellt sind und über mehr Expertise verfügen. Diese Hypothese ist nicht zu überprüfen. Gleichwohl zeigt sich nach der Auswertung der Daten ein heterogenes Bild in Bezug auf die Positionierung und Expertise zu Inklusion. So lässt sich durch die Interviews ein breites Spektrum an Positionen abbilden. Dies lässt sich unter anderem damit erklären, dass nicht nur Personen interviewt wurden, die für den Themenbereich Inklusion zuständig sind.

6.2.5 Qualitative Inhaltsanalyse der Experteninterviews

Die Experteninterviews wurden – wie die Instrumente – mit Hilfe der qualitativen Inhaltsanalyse ausgewertet. Da die Schritte des Verfahrens bereits dargestellt wurden (→ Kapitel 6.1.3), wird im Folgenden nur auf die konkrete Anwendung des Verfahrens in Bezug auf die Interviews eingegangen.

Vorbereitung bei der Extraktion für die Analyse der Interviews: Generell wurde im Kategoriensystem zwischen drei Themenbereichen unterschieden, die auf den theoretischen Vorüberlegungen basieren. Jene Themenbereiche wurden auch explizit im Leitfaden angesprochen (→ Abschnitt A im Anhang). Die verwendeten Kategorien werden im Folgenden näher erläutert.

a) **Verständnis von „Inklusion“ beziehungsweise Instrumente zu Inklusion als Ergebnis des Sensemaking-Prozesses**

In diesem Themenbereich geht es um die Frage, was die einzelnen Mitarbeiter_innen der Schulinspektion unter Inklusion verstehen. Dies kann davon abweichen, wie in den Bewertungskriterien des jeweiligen Bundeslands Inklusion verwendet wird. Bei der Auswertung dieses Teils wurde auf ähnliche Kategorien zurückgegriffen, wie bei dem zweiten Schritt der Analyse der Bewertungsinstrumente. Beide Analysen basieren auf den Dimensionen, die in Abschnitt 2.1 als Grundlage dienen, um verschiedene Inklusionsverständnisse voneinander zu unterscheiden, z. B. Zielgruppe, Begründung für Inklusion. Jedoch wurden während der Analyse der Interviews noch weitere Kategorien ergänzt. Diese dienten zum Teil dazu eine Typenbildung vorzunehmen.

b) **Rahmenbedingungen bei der Entwicklung der Instrumente**

Hier geht es darum, welche Rahmenbedingungen das Sensemaking geprägt haben. Spillane et al. (2002) unterscheiden hierbei zwischen drei Elementen: Attributen des Sensemakers, dem situativen Kontext sowie der politische Vorgabe. Hier ergaben sich einige Kategorien durch die theoretische Auseinandersetzung mit Schulinspektionen in Deutschland (→ Kapitel 3). Darüber hinaus wurden einige Kategorien induktiv während der Analyse gebildet. Als Beispiele für Kategorien sind unter anderem zu nennen:

- Element „persönliche Attribute“, z. B. Qualifikation.
- Element „situativer Kontext“, z. B. Ministerium, Forschung.
- Element „politische Vorgabe“, z. B. BRK, Schulgesetz, rechtliche Vorgaben im Allgemeinen.

c) **Bewertungssoziologische Aspekte**

In diesem Themenbereich geht es darum, wie sich die Gesprächspartner sich zu bewertungssoziologischen Aspekten positionieren. Kategorien sind: „Objektivität der Kriterien“; „Legitimation der Kriterien“; „Verständnis von guter bzw. schlechter Schule“; „Debatte zu Schulqualität“ sowie „gute Evaluation“.

Extraktion und Paraphrasierung als ein Vorgehensschritt bei der Analyse der Interviews:

Bei der Extraktion stellte sich insbesondere der Themenbereich „(b) Rahmenbedingungen bei der Entwicklung der Instrumente“ als komplex heraus. Dies hängt damit zusammen, dass zwischen verschiedenen Aspekten zu unterscheiden ist: Welche Rahmenbedingungen und Akteure waren im Umfeld der Schulinspektion von Relevanz? Welche Impulse gingen von diesen Rahmenbedingungen in Bezug auf die Deutung von Inklusion sowie auf die Entwicklung von Instrumenten aus? Wie wirken sich diese dann tatsächlich auf die Deutung von Inklusion und die Instrumente zu Inklusion aus? Des Weiteren ist zu berücksichtigen, dass in den Interviews sowohl thematisiert wurde, wie sich der Interviewpartner als Einzelperson positioniert und welche Entscheidungen auf kollektiver Ebene innerhalb der Schulinspektion getroffen wurden. In Bezug auf letzters variierten die Aussagen z. T. zwischen den Gesprächspartnern innerhalb eines Bundeslandes (nähere Ausführung folgt).

Nach der ersten Sichtung der Interviews sowie Strukturierung des Kategoriensystems wurde eine Software zur qualitativen Textanalyse dazu genutzt Textbausteine, zu markieren, um unterschiedlichen Codes hinzuzufügen. Da diese keine ausführliche Kommentierung von einzelnen Textcodierungen zugelassen hat und die komplexen Zusammenhänge somit nur begrenzt erfasst werden konnte, wurde als Ergänzung mit einem Tabellenkalkulationsprogramm gearbeitet. Die Paraphrasierung aller Codes fand daher parallel zur Extraktion statt. Zugleich wurden die einzelnen Textbausteine in der Software für die qualitative Inhaltsanalyse mit einem dreiteiligen Code in der Kommentierung benannt: z. B. „1A_3“. Die erste Ziffer steht für die Nummer des Interviews, der nachfolgende Buchstabe steht für das Bundesland und die letzte Ziffer steht für die Nummer des Textbausteins innerhalb des Interviews. Bei der Darstellung der Ergebnisse wird davon abgesehen, den Buchstaben des Bundeslandes aufzuführen, um die Anonymität der Fälle zu gewährleisten.

Auswertung bei der Analyse der Interviews: In der Auswertung werden drei Bereiche des Kategoriensystems (→ Abschnitt 6.2.5) unterschieden. Das Vorgehen **beim ersten Themenbereich** „(a) Verständnis von Inklusion bzw. Bewertungsinstrumenten zu Inklusion“ wies Ähnlichkeiten zu der Analyse der Instrumente auf. Jedoch zeigte sich, dass andere Kategorien von Relevanz waren, um einzelne Typen voneinander zu unterscheiden. Beim **zweiten Themenbereich** „(b) Rahmenbedingungen bei der Entwicklung der Instrumente (Sensemaking)“ kristallisierte sich heraus, dass in den einzelnen Bundesländern zwar ähnliche Rahmenbedingungen prägend waren. Es traten jedoch Unterschiede darin auf, welche Impulse davon ausgingen, z. B. dass die Kultusministerien in den einzelnen Bundesländern voneinander abweichende Positionen vertraten. Des Weiteren unterschieden sich die Instrumenteentwickler_innen auch darin, wie sie auf die jeweiligen Impulse – auch wenn es die gleichen waren – reagierten. Dies wurde bei der Darstellung der Ergebnisse insofern berücksichtigt, als im ersten Schritt aufgeführt wurde, welche Rahmenbedingungen und Akteure generell eine Rolle gespielt haben. Durch die Beschreibung von exemplarischen Einzelfällen wird dargestellt, wie die einzelnen Elemente miteinander interagieren und sich daraus unterschiedliche Implikationen ergeben. Bei der Auswertung des **dritten Themenbereichs** „(c) Bewertungssoziologische Aspekte“ zeigte sich, dass die Positionierung zu jenen Aspekten beim Sensemaking-Ansatz dem Element „persönliche Attribute des Sensemakers“ zugeordnet werden können.

6.2.6 Grenzen der Datenanalyse

Im Folgenden wird dargestellt, welche Grenzen mit dem gewählten Datenerhebungsformat, dem Vorgehen sowie dem Datenzugang einhergehen.

Prozess und Kausalität: Durch die gewählten Datenerhebungsformate (Dokumentenanalyse sowie Experteninterviews) ist es nur begrenzt möglich, die Entwicklung der Instrumente als Prozess abzubilden. Hierfür wären Informationen über Beginn und Ende des Prozesses notwendig. So fehlen Informationen darüber, welche Wissensbestände und Überzeugungen vorhanden waren, als die Instrumenteentwickler_innen angefangen haben sich mit Inklusion auseinanderzusetzen. So hatte zum Zeitpunkt der Datenerhebung der Prozess der Auseinandersetzung zu Inklusion bei den Schulinspektionen bereits begonnen und es ist keine Dokumentation über diesen Zeitraum vorhanden. Demnach lässt sich die Art der Auseinandersetzung mit Inklusion

bzw. inklusiven Bewertungsinstrumenten primär nur als Querschnitt herausarbeiten. Dennoch gibt es in den Interviews subjektive Einschätzungen über den Verlauf des Prozesses. Darüber hinaus erhebt der Sensemaking-Ansatz nicht den Anspruch, Aussagen über Kausalitäten zu treffen (→ Kapitel 5). Demnach gibt es keine kausalen Mechanismen, die sich linear abbilden lassen. Dies hängt unter anderem auch damit zusammen, dass Sensemaking ein zirkulärer Prozess ist, der sich dadurch kennzeichnet, dass einzelne Rahmenbedingungen miteinander interagieren. Dies kann bei der Auswertung zu Problemen führen, die umgangssprachlich als „Henne-Ei-Problem“ bezeichnet werden. So kann das Verständnis von Inklusion sowohl als das Ergebnis eines Deutungsprozesses verstanden werden, aber auch als eine Rahmenbedingung, der die Entwicklung von Bewertungsinstrumenten zu Inklusion prägt.

Fehlende Verknüpfung zwischen Interviews und Instrumenten: Die analysierten Bewertungsinstrumente sind frei zugänglich. Daher können die Bundesländer bei der Analyse der Instrumente namentlich genannt werden. Dagegen wurde den Interviewpartnern zugesichert, dass die Bundesländer, in denen sie tätig sind, nicht genannt werden. In Bezug auf die Arbeit bedeutet dies, dass zwischen der Analyse der Instrumente und der Interviews keine direkten Verknüpfungen auf Bundeslandebene hergestellt werden können. Beispielsweise lässt sich nicht explizit darstellen, warum Bundesland X bestimmte Kriterien nicht verwendet.

Fehlende Informationen zum kollektiven Aushandlungsprozess: Im Rahmen der Arbeit geht es zum einen um die Deutung von Inklusion bzw. inklusionsorientierten Bewertungsinstrumenten und zum anderen um die Rahmenbedingungen, die die Auseinandersetzung mit Inklusion geprägt haben. Wie bereits oben dargestellt, ist in der Analyse zwischen verschiedenen Ebenen zu unterscheiden:

- die individuelle Ebene (Instrumenteentwickler_innen);
- die kollektive Ebene, d.h. die Aushandlungsprozesse innerhalb eines Bundeslandes;
- die generisch-subjektive Ebene in Form der Messinstrumente.

Um erklären zu können, wie die Bewertungsinstrumente als Ergebnis von kollektiven Aushandlungsprozessen zustande gekommen sind, wäre es notwendig gewesen teilnehmende Beobachtungen durchzuführen. Dies hätte es ermöglicht, Interaktionen zwischen diversen Akteuren zu beobachten und einen Einblick in den Prozess zu erhalten. Die Entwicklung der Instrumente war jedoch zu Beginn der Arbeit in den meisten Bundesländern bereits abgeschlossen. Darüber hinaus wäre es schwer machbar gewesen, teilnehmende Beobachtungen mit mehreren Bundesländern zu realisieren, weil sich die Entwicklung von Instrumenten über mehrere Monate bzw. Jahre hinzieht und eine Vielzahl von Akteuren involviert ist.

Stattdessen wurde Experteninterviews als Erhebungsformat gewählt. Hierbei standen die individuellen Sensemaking-Prozesse im Fokus der Interviews. Gleichwohl werden auch kollektive Aushandlungsprozesse thematisiert, aber hierbei aus der Perspektive der einzelnen Interviewpersonen. Demnach dienen die Interviews mit einzelnen Instrumenteentwickler_inen lediglich als Annäherung daran, welche Rahmenbedingungen die Entwicklung auf kollektiver Ebene geprägt haben. Es kann daher sein, dass die für das Individuum relevanten Rahmenbedingungen nur Teilaspekte widerspiegeln bzw. dass bestimmte Rahmenbedingungen nur für die interviewten Personen eine große Rolle gespielt haben, aber auf der Ebene des kollektiven Aushandlungsprozesses keine Relevanz mehr hatten. Nichtsdestotrotz sind die herausgearbeiteten Rahmenbedingungen von großer Relevanz, da durch diese nachvollziehbar wird, wie es zu den

jeweiligen Instrumenten gekommen ist. Des Weiteren zeigt sich in der Analyse der Interviews, dass es auf Fallebene viele Überschneidungen bzw. ähnliche Einschätzungen über relevante Akteure und Rahmenbedingungen bei der Entwicklung der Instrumente gibt.

Anwendung der Bewertungsinstrumente: Die Arbeit kann keine Aussagen darüber treffen, wie die Bewertungsinstrumente in der konkreten Praxis angewandt werden. So findet während der Inspektionen erneut eine Deutung der Instrumente durch die Inspektor_innen statt (vgl. Lambrecht 2013, S. 238; Sowada und Dederich 2014; Lamont und Huutoniemi 2011). Diese ist von verschiedenen Faktoren geprägt, z. B. Rahmenbedingungen der Schule, eigene Erwartungen an Schulen. Dieser Aspekt würde den Rahmen der Arbeit sprengen, wird jedoch in den Arbeiten von Moritz Sowada und Fanny Oehme ausführlich bearbeitet.

6.3 Forschungsethik und Datenschutz

Bei der Durchführung des Forschungsprojekts wurden ethische Implikationen immer wieder reflektiert und im Handeln berücksichtigt. Dies spielte insbesondere bei den Interviews eine Rolle. Es wird im Folgenden dargestellt, welche Maßnahmen in diesem Bereich getroffen wurden.

6.3.1 Informierte Einwilligung

In der Kontaktaufnahme wurden die Interviewpartner über die Fragestellung (vgl. Gebel et al. 2015) sowie über das Thema des Interviews informiert. Darüber hinaus erhielten sie Informationen über die Datenaufbewahrung sowie über die Autorin der Arbeit. Vor Interviewbeginn wurden die Interviewten mündlich nochmal über das Thema des Forschungsprojekts informiert. Des Weiteren wurden sie darauf hingewiesen, dass das Interview freiwillig sei und mit einem Aufnahmegerät aufgezeichnet wird. Zusätzlich erhielten sie erneut Informationen zur Datenaufbewahrung und zur Datennutzung. Die Interviewten stimmten mündlich zu, dass sie mit dem weiteren Verfahren einverstanden sind. Von einer schriftlichen Einwilligung wurde abgesehen, um möglichst nah an einer natürlichen Gesprächssituation zu bleiben. Im Rahmen des Interviews wurden keine sensiblen Daten, z. B. Herkunft, religiöse Überzeugungen usw. erhoben. Des Weiteren wurde den Interviewten zugesichert, dass sie über die Ergebnisse der Forschung informiert werden. Zusätzlich zu den Interviews hat die Autorin bei einer Schulinspektion hospitiert und an bundesländerübergreifenden Arbeitstreffen teilgenommen. Die dort erworbenen Informationen sind nicht Teil der Datenanalyse, spielen jedoch als Hintergrundwissen eine Rolle. Den Teilnehmer_innen der Arbeitstreffen war von Anfang an bekannt, dass die Autorin dieser Arbeit dort in ihrer Rolle als Forscherin teilnimmt und welches Forschungsthema bearbeitet wird.

6.3.2 Datenschutz

Es wurden verschiedene Maßnahmen getroffen, um die Sicherheit der Interviewdaten zu gewährleisten. Bei den Interviews wurde den Interviewten zugesichert, dass die Namen der Interviewpartner, die Organisationseinheiten sowie die Bundesländer anonym bleiben. Daher wurde auch die Liste mit den Namen der Interviewpartner und deren Verknüpfung mit den verwendeten CaseIDs von den Transkriptionen getrennt aufbewahrt, damit kein Personenbezug

herstellbar ist. Die Audiofiles wurden auf dem Cloud-Server des Wissenschaftszentrums Berlin für Sozialforschung aufbewahrt. Diese waren nur mit den Nutzerdaten der Autorin zugänglich. Bei der Transkription wurde darauf geachtet, die Daten zu anonymisieren, indem nicht nur die Namen der Interviewten, sondern auch Namen der Organisation, des Bundeslands, usw. entfernt wurden. Des Weiteren zeigte sich bei der Analyse, dass bestimmte Fälle aufgrund spezifischer Rahmenbedingungen oder Verwendung von Dialekten trotz Anonymisierung möglicherweise noch identifiziert werden konnten. Daher wurden in der Verschriftlichung Originalzitate abgeändert, um mögliche Identifizierungen der Interviewpartner zu vermeiden. Dies wurde bei den jeweiligen Originalziten gekennzeichnet.

Es folgt nun die Darstellung der empirischen Analyse: Dabei geht es in Kapitel 7 um die Deutungen von Inklusion in den Instrumenten und bei den Instrumenteentwickler_innen und in Kapitel 8 um die Entwicklung der Instrumente.

7 Deutungen von Inklusion

Im Rahmen des folgenden Kapitels geht es um Deutungen von Inklusion in den Instrumenten der Schulinspektionen sowie bei den interviewten Instrumenteentwickler_innen. Wie in Kapitel 6 bereits dargestellt, werden verschiedene Verfahren miteinander kombiniert. Daher besteht das Kapitel aus den folgenden Abschnitten:

- **Abschnitt 7.1: Implizite Berücksichtigung von Inklusion als normative Forderung in den Instrumenten:** In diesem Abschnitt geht es darum, ob Inklusion als normative Forderung in den Instrumenten berücksichtigt wird. Als Orientierung dient das Verständnis von Inklusion, welches in Abschnitt 2.2 herausgearbeitet wurde. Bei der Analyse sind alle verfügbaren Instrumente Gegenstand der Analyse und nicht nur diejenigen Textbausteine, in denen Inklusion explizit als Begriff verwendet wird. Bei der folgenden Darstellung wird zum einen betrachtet, welche Formen der Operationalisierung gewählt wurden, und zum anderen, inwiefern sich die Bundesländer in der Berücksichtigung von Inklusion unterscheiden. Diese Analyse lässt Aussagen darüber zu, inwiefern Inklusion als politische Vorgabe implementiert ist – unabhängig davon, ob dies explizit mit dem Inklusionsbegriff verknüpft wird.
- **Abschnitt 7.2: Explizite Verwendung des Inklusionsbegriffs in den Instrumenten:** In diesem Teil der Analyse geht es darum, wie Inklusion als Begriff in den Instrumenten verwendet wird. Damit wird gezeigt, welche Deutungen von Inklusion in den Instrumenten existieren und inwiefern sich die Bundesländer in Bezug auf ihre Deutungen von Inklusion voneinander unterscheiden. Hierfür wird nur auf Textbausteine zurückgegriffen, in denen Inklusion explizit als Begriff verwendet wird bzw. die Instrumente als inklusionsorientiert gedeutet werden können. Diese Analyse lässt Aussagen darüber zu, wie Inklusion von den Schulinspektionen in den einzelnen Bundesländern gedeutet wurde.
- **Abschnitt 7.3: Deutungen von Inklusion von Seiten der Instrumenteentwickler_innen durch die Analyse der Interviews:** In diesem Teil wird dargestellt, welche Verständnisse von Inklusion bei den einzelnen Instrumenteentwickler_innen vertreten sind. Deren Deutungen stellen das Ergebnis von individuellen Sensemaking-Prozessen dar. Sie dienen als Grundlage, um im folgenden Kapitel herauszuarbeiten, welche Rahmenbedingungen die Deutung geprägt haben und somit zu den jeweiligen Deutungen und Messinstrumenten geführt haben. Abschließend wird dargestellt, inwiefern sich die Deutungen von Inklusion in den Instrumenten und bei den Instrumenteentwickler_innen unterscheiden. In diesem Zusammenhang geben die Interviews erste Hinweise darüber, wie und warum sich bestimmte Deutungen der Interviewpartner in den Instrumenten durchsetzen konnten bzw. nicht durchgesetzt haben.

Bei der Betrachtung der Instrumente sind die folgenden Aspekte zu berücksichtigen:

- **Unterscheidung zwischen Wahl- und Pflichtbereich:** In manchen Bundesländern, z. B. Sachsen-Anhalt, wird zwischen Wahl- und Pflichtbereich unterschieden. Der Pflichtbereich ist für alle Schulen obligatorisch. Beim Wahlbereich können die Schulen auswählen, in welchen Themenbereichen sie evaluiert werden möchten. Diese Unterscheidung zwischen Wahl- und Pflichtbereich kann so gelesen werden, dass bestimmte Themenbereiche von Seiten der Schulinspektionen in Bezug auf Schulqualität für relevanter erachtet werden als andere bzw. dass den Schulen Mitspracherecht eingeräumt werden soll. Jedoch kann die Unterscheidung

zwischen Wahl- und Pflichtbereich auch damit begründet werden, dass bestimmte Wahlbereiche nicht auf alle Schulen zutreffen, z. B. ist die Evaluation des Ganztagsbetriebs nur sinnvoll, wenn dieser an einer Schule existiert. Im Folgenden wird bei der Quellenangabe gekennzeichnet, wenn ein Indikator aus einem Teil des „Wahlbereichs“ ist bzw. es wird auch thematisiert, wenn Inklusion als Ganzes einen eigenen Qualitätsbereich im Wahlbereich darstellt. Des Weiteren gibt es auch den Fall, dass eine Inspektion keinen Wahlbereich hat, aber trotzdem bestimmte Instrumente nur zur Anwendung kommen, wenn diese für die schulspezifische Situation von Relevanz sind, z. B. in Baden-Württemberg (nähere Ausführung folgt).

- **Analyse der Instrumente sowie Berücksichtigung der Ergebnisdarstellung:** Bei der Analyse der Instrumente in Bezug auf Inklusion ist auch von Relevanz, wie die Daten, die mithilfe der Instrumente erhoben werden, in den Schulinspektionsberichten aufbereitet und dargestellt werden (→ Kapitel 4). Hierfür wird exemplarisch auf Inspektionsberichte zurückgegriffen, falls diese verfügbar waren.
- **Unterscheidung zwischen evaluativem und deskriptivem Bereich:** In Bezug auf die Darstellung der Daten unterscheiden manche Bundesländer zwischen einem evaluativen und einem deskriptiven Bereich. Im deskriptiven Bereich werden die erhobenen Daten dargestellt, aber es findet keine Wertung von Seiten der Schulinspektionen statt, z. B. im Fall des Unterrichtsbeobachtungsbogens in Niedersachsen (UBB_NI).

7.1 Berücksichtigung von Konzepten in Bezug auf Inklusion

Im Folgenden wird näher betrachtet, ob und wie die Ziele von Inklusion („individuell bestmögliche Qualifizierung“, „politische Partizipation“, „Erfahren von intersubjektiver Anerkennung“) sowie die Differenzlinien in den Messinstrumenten operationalisiert werden.

7.1.1 Ziel von Inklusion: Individuell bestmögliche Qualifizierung

Um abzubilden, wie Differenzierung im Lernprozess in den Instrumenten operationalisiert wird und inwiefern sich die Bundesländer voneinander unterscheiden, bietet es sich an drei Aspekte näher zu betrachten: Arten der Differenzierung, wahrgenommene Heterogenität einer Lerngruppe in den Messinstrumenten sowie Formen der Outputmessung.

Arten der Differenzierung: In Bezug auf den Lernprozess werden verschiedene Arten der Differenzierung von Seiten der Schulinspektionen abgefragt:

- Kennen und Berücksichtigung der unterschiedlichen Lernstände, z. B. „Die Lehrkraft informiert sich über individuelle Lernschwierigkeiten“ (UBB_BY).
- Ansprechen von unterschiedlichen Lerntypen durch das Anbieten von unterschiedlichen Aufgabentypen, z. B. „Die Lehrkraft bietet Wahlmöglichkeiten oder erteilt differenzierte Arbeitsaufträge bezüglich der Arbeitsform, Informationsquelle oder der Ergebnisdarstellung“ (UBB_HE).
- Qualitative Differenzierung, z. B. Aufgaben mit unterschiedlichen Schwierigkeitsgraden, z. B. „Die Schülerinnen und Schüler erhalten differenzierte Lernangebote durch verschiedene Aufgabenschwierigkeiten“ (UBB_BW).

- Quantitative Differenzierung durch Variation innerhalb der Lerngruppe in Bezug auf die Anzahl der Aufgaben bzw. die gegebene Zeit, z. B.
 - „Es gibt Differenzierung nach Umfang bzw. Zeit“ (UBB_NRW).
 - „Für schnell arbeitende Schüler steht zusätzliches Arbeitsmaterial bereit“ (UBB_BY).
- Leistungsrückmeldung, z. B.
 - „Es gibt Formen der differenzierten Leistungsrückmeldung“ (UBB_BE).
 - „Die Schülerinnen und Schüler erhalten differenzierte Leistungsrückmeldungen“ (UBB_HH).
- Zusatzangebote, z. B. „Die zusätzliche Förderung ist auf das bestmögliche Lernen aller Schülerinnen und Schüler ausgerichtet“ (ÜI_HH).
- Wahlmöglichkeiten in Bezug auf Themen und Inhalte, z. B. „Die Lehrkraft bietet Wahlmöglichkeiten oder erteilt differenzierte Arbeitsaufträge bezüglich der Themen / Inhalte“ (UBB_HE).
- Förderplan, z. B. „Förderplangespräche mit den betreffenden Eltern“ (LFB_HE).

In den Instrumenten sowie in den „Merkmale von guter Schule“ (ÜI) werden in allen Bundesländern verschiedene Aspekte der Differenzierung aufgeführt. Die Länder unterscheiden sich jedoch darin, welche und wie viele Arten der Differenzierung, z. B. bei der Leistungsrückmeldung, bei den Aufgabenformaten usw., aufgeführt werden. Da es zwischen den Bundesländern variiert, welche Dokumente zugänglich sind und welches Format die Instrumente aufweisen, ist der Vergleich der Bundesländer über die Anzahl der Nennungen nicht aussagekräftig und wird daher nicht realisiert.

Wahrgenommene Heterogenität innerhalb der Lerngruppe: Im vorherigen Abschnitt ist deutlich geworden, dass es verschiedene Möglichkeiten gibt, die Heterogenität der Lerngruppe in der Unterrichtsgestaltung zu berücksichtigen. Hierbei ist jedoch zu beachten, wie die Lerngruppe bei der jeweiligen Differenzierungsform unterteilt wird. Daher wird im Folgenden dargestellt, auf welche Art und Weise in den Instrumenten die Heterogenität der Lerngruppe beschrieben wird. So kristallisierten sich bei der Analyse drei Formen heraus:

- a) Unterteilung der Lerngruppe in kleinere Gruppen ohne genauere Angaben, auf welcher Basis die Unterteilung stattfindet, z. B. „Es gibt ein vorbereitetes, differenziertes Unterrichtsangebot in Bezug auf Umfang und Zeit“ (UBB_NI).
- b) Unterteilung der Lerngruppe auf der Grundlage einer Differenzlinie, z. B. „In meiner Schule werden leistungsstarke Schüler/innen besonders gefördert“ (LFB_BE) oder „Leistungsschwächere Schüler/innen erhalten in meiner Schule besondere Unterstützung“ (LFB_BE).
- c) Betrachtung auf Individualebene, z. B. „Die Lehrkraft informiert sich über den individuellen Lernfortschritt“ (UBB_BY) oder „Einzelne Schüler werden gezielt gefördert“ (ÜI_BLI).

In Bezug auf die erste Unterscheidungsform (siehe a) fehlt es an Informationen darüber, anhand welcher Differenzlinie(n) Unterteilungen innerhalb der Lerngruppe vorgenommen werden. Es bleibt offen, ob in diesem Fall die Schüler_innen in ihrer Individualität wahrgenommen werden (wie bei (c) Betrachtung auf Individualebene). Es ist auch möglich, dass auf der Grundlage einer Differenzlinie zwischen zwei Gruppen unterschieden wird, z. B. leistungsstark/leistungsschwach. Dies würde dann der zweiten Unterscheidungsform „(b) Unterteilung der Lerngruppe auf der Grundlage einer Differenzlinie“ entsprechen. Ausgehend von dem hier verwendeten Verständnis von Inklusion würde diese Form der Operationalisierung nicht

als Instrument zur Erfassung von Inklusion verstanden werden. Danach spielen bei einem Individuum stets diverse Differenzlinien in ihrer Verschränkung eine Rolle.

Es zeigt sich, dass in den einzelnen Bundesländern immer auf verschiedene Unterteilungsformen in Bezug auf die Lerngruppe zurückgegriffen werden, das heißt die oben beschriebenen Formen werden miteinander kombiniert. Demnach gibt es kein Bundesland, welches bei den Items nur auf der Individualebene oder auf der Gruppenebene differenziert. Bundesländerübergreifend zeigt sich jedoch, dass bei den meisten Items die erste Unterscheidungsform, (a) Unterteilung der Lerngruppe ohne genauere Angabe, gewählt wird, d. h. es fehlen Angaben darüber, welche Form der Unterteilung innerhalb der Lerngruppe stattfindet. Als bundesländerübergreifendes Muster zeigt sich, dass die Differenzierung auf individueller Ebene insbesondere in Bezug auf das „Kennen und Berücksichtigen“ des Lernstandes eine Rolle spielt (z. B. UBB_BY; UBB_BB; UBB_HH; UBB_ST; ÜI_BLI).

Im Fall von *Schleswig-Holstein* fehlt es an genaueren Angaben im Unterrichtsbeobachtungsbogen, da dieser qualitativ angelegt ist und dementsprechend Stichwörter enthält, die in einem offenen Textfeld kommentiert werden können. Das Thema „Unterteilung der Lerngruppe“ wird insofern angesprochen, als es das Stichwort „Umgang mit Heterogenität“ gibt und Schulinspektor_innen dazu aufgefordert werden, offene Angaben zu diesem Aspekt zu machen. Es gibt jedoch keine Informationen darüber, welche Formen der Unterteilung innerhalb der Lerngruppe angestrebt werden sollte.

Bei der zweiten Unterscheidungsform, der Unterteilung der Lerngruppe auf der Grundlage einer Differenzlinie, wird stets auf die gleiche Differenzlinie zurückgegriffen: die Leistungsfähigkeit der Schüler_innen. Demnach spielen andere Differenzlinien bei der Unterteilung der Lerngruppe im Hinblick auf Lernprozesse keine Rolle. Bei der Beschreibung von Leistungsfähigkeit unterscheiden sich die Bundesländer darin, welche Adjektive verwendet werden. Dies wird hier am Beispiel des Schülerfragebogens (SFB) illustriert (→ Tab 7.1).

Hierbei zeigt sich als Muster, dass überwiegend auf dichotome Unterscheidungen zurückgegriffen wird. Die Fokussierung auf dichotome Kategorien ist aus Inklusionsperspektive kritisch zu betrachten, da damit stabile Rollenzuschreibungen einhergehen. Dies ist insbesondere problematisch, wenn diese mit negativen Zuschreibungen verknüpft sind (nähere Ausführung folgt). Eine Ausnahme stellen Brandenburg, Hamburg und Hessen dar, da in vielen Fällen keine dichotome Unterscheidung in Bezug auf Leistungsfähigkeit gewählt wird, sondern die Schüler_innen auf einem Spektrum verortet werden können, z. B. „Meine Lehrerinnen und Lehrer geben uns unterschiedlich viele Aufgaben, je nachdem, wie viel der Einzelne schafft“ (SFB_HE).

Die dritte Unterscheidungsform, die Betrachtung der Individualebene, wird im Rahmen der Schülerfragebögen beispielsweise so operationalisiert, dass die einzelnen Schüler_innen gefragt werden, inwiefern der Einzelne im Unterricht berücksichtigt wird, z. B. „Die Lehrkräfte kennen die Stärken und Schwächen der einzelnen Schülerinnen und Schüler“ (SFB_ST; siehe auch SFB_HE). Des Weiteren nutzen viele Bundesländer das Format des Schülerfragebogens, um die Schüler_innen um Selbsteinschätzungen zu bitten, z. B. „Ich fühle mich in der Schule zur Zeit überfordert“ (SFB_SL; siehe auch SFB_BY; SFB_BE; SFB_BB; SFB_HE; SFB_ST; SFB_SH). Dieses Format ermöglicht es, Selbsteinschätzungen auf individueller Ebene zu erfassen und gibt somit erste Hinweise in Bezug auf das Ziel „individuell bestmögliche Qualifizierung“. In diesem Zusammenhang ist jedoch zu berücksichtigen, wie die Werte von Seiten der Schulinspektion in den Berichten dargestellt werden. Beispielsweise gibt es im Fall von Brandenburg (IB_BB; in Bezug auf Darstellung siehe auch IB_BE; IB_ST) das Item „Wenn ich zusätzliche Hilfe brauche,

Tab. 7.1: Operationalisierung von Leistung

Verwendete Adjektive	Konkrete Operationalisierung
Leistungsstark/ -schwach	„leistungsstärkeren/-schwächeren“ (SFB_BB); schwächeren Schülerinnen (SFB_ST)
Langsam oder Lernschwierigkeiten/schnell	für langsamere/schnellere Schüler_innen (SFB_SL); „die Lernschwierigkeiten haben und oder an diejenigen, die schneller vorankommen“ (SFB_HH); „wenn man besonders gut und schnell lernt“ (SFB_HH); „SuS, die Lernschwierigkeiten haben und/oder SuS, die schneller vorankommen“ (SFB_SH).
Gut /schwer fallen	„besonders gute Schülerinnen und Schüler“ (SFB_HE); „Wenn Kinder etwas besonders gut können“ (SFB_BE); „denen das Lernen schwer fällt“ (SFB_HE); „Wenn Kindern im Unterricht etwas schwerer fällt“ (SFB_BE);
Keine dichotome Leistungsunterscheidung	„je nach unserem Können“ (SFB_BB); „Gruppen (...), die nach unseren Fähigkeiten eingeteilt sind“ (SFB_HH); „Meine Lehrerinnen und Lehrer geben uns unterschiedlich viele Aufgaben, je nachdem wie viele der Einzelne schafft“ (SFB_HE)

bekomme ich sie von meiner Lehrerin oder meinem Lehrer“. Die Schüler_innen treffen Aussagen auf einer Skala von 1 („stimme nicht zu“) bis 4 („stimme voll und ganz zu“). Bei der Auswertung im betrachteten Inspektionsbericht von Brandenburg (IB_BB) wurde der Durchschnitt errechnet. Es fehlen jedoch Aussagen zur Verteilung der Werte. So kann es sein, dass eine niedrige Standardabweichung besteht und alle Schüler_innen ähnliche Werte vergeben haben. Es besteht theoretisch auch die Möglichkeit, dass ein geringer Teil der gesamten Schülerschaft den Wert „1“ angegeben hat. Durch die Angabe des Durchschnittswerts fehlt die Information, ob dies der Fall ist und wenn ja, wie viele Schüler_innen niedrige Werte angegeben haben. Dies ist aus Inklusionsperspektive problematisch, denn es geht darum, dass alle Schüler_innen individuell bestmöglich gefördert werden und somit auch die Situation von Schüler_innen, die innerhalb einer Lerngruppe zahlenmäßig gering vertreten sind. So kann es beispielsweise sein, dass eine bestimmte Gruppe von Schüler_innen, z. B. diejenigen mit sonderpädagogischem Förderbedarf, im Schülerfragebogen Einschätzungen darüber abgibt, dass sie nicht bestmöglich unterstützt wird. Da diese Gruppe jedoch im Vergleich zu Schüler_innen ohne sonderpädagogischen Förderbedarf innerhalb der Lerngruppe in der Unterzahl ist, würde ihre Situation, falls diese abweichend vom Durchschnitt ist, nicht sichtbar werden. Demnach bietet das Instrument zwar das Potenzial, Aussagen darüber zu treffen, inwiefern alle Schüler_innen bestmöglich qualifiziert werden. Indem jedoch bei der Darstellung der Ergebnisse bestimmte Attribute der Schüler_innen, z. B. Geschlecht oder sonderpädagogischer Förderbedarf, nicht mit dem Antwortverhalten verknüpft werden und die Verteilung der Antworten nicht dargestellt wird, lassen sich nur begrenzt Aussagen darüber treffen, inwiefern alle Schüler_innen bestmöglich qualifiziert werden. Beim Elternfragebogen zeigen sich im Vergleich zum Schülerfragebogen

ähnliche Formen der Operationalisierung und damit einhergehende Implikationen. So wird beispielsweise die Individualebene hier folgendermaßen abgefragt: „Ich habe den Eindruck, dass mein Kind im Unterricht angemessen gefördert wird“ (EFB_BB).

Eine ausführlichere Aufschlüsselung der Daten zeigt sich im betrachteten Inspektionsbericht aus Schleswig-Holstein (IB_SH). Beim Indikator „Meine Lehrerin/ mein Lehrer sagt mir, was ich tun kann, um mich zu verbessern“ werden die einzelnen Werte aufgeführt. Im Fall der Beispielschule gaben 27 Prozent an „in fast keiner Stunde“; 45,4 Prozent „in einigen Stunden“; 17,6 Prozent „in den meisten Stunden“ und 6,3 Prozent der Schüler_innen gaben an, dass sie „in jeder Stunde“ individuelle Rückmeldung darüber erhalten, wie sie sich verbessern können. Durch diese Aufschlüsselung der Daten werden somit Informationen zugänglich, wie die Werte innerhalb der Schülerschaft verteilt sind. Es gibt jedoch auch in diesem Fall keine Informationen darüber, ob mögliche Erklärungen für die unterschiedlichen Werte darin liegen, dass systematisch bestimmte Schülergruppen nicht ausreichend gefördert werden.

Zusammenfassend lässt sich beobachten, dass in dem Dokumententyp „Merkmale von guter Schule“ (ÜI) sowie im UBB der Aspekt heterogene Lerngruppe häufig angesprochen wird, aber nähere Angaben dazu fehlen, wie die Lerngruppe unterteilt bzw. betrachtet werden soll. Ausgehend von dem in dieser Arbeit verwendeten Verständnis von Inklusion würden diese Instrumente nicht als „inklusionsorientiert“ verstanden werden, wenn sie sich stets auf gleichbleibende dichotome Unterscheidungen beziehen, z. B. Schüler_innen mit und ohne sonderpädagogischen Förderbedarf. Darüber hinaus bieten die Eltern- und Schülerfragebögen die Möglichkeit, Einschätzungen auf Individualebene zu erfassen. Hierbei ist zu berücksichtigen, wie die Abfragen im nächsten Schritt für die Rezipienten der Daten aufbereitet werden. Wenn nur Durchschnittswerte dargestellt werden, kann es sein, dass die spezifische Situation von bestimmten Schülergruppen durch die Datendarstellung nicht sichtbar wird.

Formen der Outputmessung: Das hier verwendete Verständnis von Inklusion (→ Kapitel 2) grenzt sich von der Annahme ab, dass es eine „normale“ Lernkurve gibt. Solche Vorstellungen spiegeln sich in schulübergreifenden Lernstandserhebungen wie den Vergleichsarbeiten wider. Wie bereits dargestellt in Abschnitt 2, können solche Operationalisierungen aus Inklusionsperspektive problematisch werden. So steht stattdessen im Fokus, inwiefern die einzelnen Schüler_innen individuell bestmögliche Lernfortschritte gemacht haben. Dieser Aspekt kann durch Querschnittsmessungen nicht abgebildet werden.

In diesem Zusammenhang stellt sich die Frage, in welcher Weise die Ergebnisse von Lernstandserhebungen bei den Schulinspektionen eine Rolle spielen. Jener Aspekt wird sowohl in den „Merkmalen von guter Schule“ sowie im „Lehrerfragebogen“ thematisiert.

Ergebnisse der Lernstandserhebungen als Indikator für Schulqualität: In einzelnen Bundesländern stellen die Ergebnisse von schulübergreifenden Lernstandserhebungen einen möglichen Output dar, um die Qualität von Schule zu messen. Hierbei ist zu berücksichtigen, dass diese Daten in vielen Fällen von Seiten der Schulinspektionen nicht bewertet werden (ÜI_BY; ÜI_BB; IB_HH; ÜI_NRW; ÜI_ST), d. h. die Ergebnisse werden im Bericht abgebildet, aber es findet keine Einschätzung von Seiten der Schulinspektionen statt, wie diese Ergebnisse zu bewerten sind.

Umgang der Schule mit den Werten: In vielen Bundesländern geht es nicht um die Frage, wie Schulen in den Vergleichsarbeiten abschneiden, sondern eher um die Frage, wie sie damit umgehen, z. B. durch folgende Indikatoren „Lernstandserhebungen und ggf. weitere stan-

dardisierte Schulleitungstests werden systematisch ausgewertet“ (ÜI_BW; siehe auch ÜI_BE; ÜI_NI; ÜI_NRW; LFB_BY; LFB_HE; LFB_SL; ÜI_HH). Nach dieser Form der Operationalisierung kennzeichnet eine gute Schule demnach nicht, dass die Schüler_innen im Querschnitt möglichst hohe Kompetenzwerte erzielen, sondern dass die Schule die Ergebnisse nutzt, um ihr eigenes Handeln zu reflektieren. Dahinter steht ein Verständnis von guter Schule, wonach sich die Schule als lernende Organisation versteht. Demnach verlieren die tatsächlichen Ergebnisse der jeweiligen Schule in den Lernstandserhebungen an Relevanz. Diese Form der Thematisierung impliziert, dass die Ergebnisse von Lernstandserhebungen zwar nicht von Seiten der Schulinspektionen, aber von Seiten der Schulen dazu genutzt werden können, um Aussagen über Schulqualität zu treffen.

Zusammenfassend lässt sich beobachten, dass in allen Bundesländern, in denen entweder die „Merkmale von guter Schule“ oder der Lehrerfragebogen zugänglich waren, Lernstandserhebungen thematisiert wurden. Jedoch unterscheiden sich die Bundesländer darin, wie sie diese berücksichtigen, z. B. werden die Ergebnisse häufig von Seiten der Bundesländer nicht bewertet bzw. sollen die Daten von Seiten der Schule für interne Schulentwicklungsmaßnahmen genutzt werden. Ausgehend von dem hier verwendeten Verständnis von Inklusion sind Daten aus Lernstandserhebungen nicht weiterführend und verfügen über das Potenzial, Schulen in Bezug auf die Aufnahme von Schüler_innen falsche Signale zu setzen.

In diesem Zusammenhang stellt sich die Frage, wie individuell bestmögliche Qualifizierung aller Schüler_innen als ein Ziel von Inklusion operationalisiert werden könnte. Ein mögliches Vorgehen ist die Befragung von Schüler_innen, ob diese den Eindruck haben, dass sie ausreichend gefordert werden (Bayern, Baden-Württemberg, Schleswig-Holstein). So wird beispielsweise im Schülerfragebogen von *Bayern* abgefragt: „Es kommt manchmal vor, dass ich auf andere warten muss, wenn ich mit meiner Aufgabe fertig bin“ (SFB_BY; auch ÜI_BW). Im Fall von *Sachsen-Anhalt* wird dieser Aspekt durch negativ formulierte Items abgefragt: „Der Unterricht ist so schwer, dass ich nicht mitkomme“ (SFB_ST) sowie „Ich fühle mich in der Schule überfordert“ (SFB_ST; siehe auch SFB_SL). In diesem Zusammenhang tritt erneut die Frage auf, wie die Ergebnisse dieser Werte dargestellt werden. So gibt es beispielsweise im Fall von *Baden-Württemberg* (IB_BW) – wie oben bereits in Bezug auf Brandenburg thematisiert – keine Aufschlüsselung darüber, welche Werte die einzelnen Schüler_innen angegeben haben, sondern nur den Durchschnittswert. Des Weiteren zeigt sich als Muster, dass der Grad an wahrgenommener Überforderung abgefragt wird und es sich nicht um eine Einschätzung der Schüler_innen darüber handelt, inwiefern sie als Individuen bestmöglich gefördert werden.

7.1.2 Ziel von Inklusion: Politische Partizipation aller

Als weiteres Ziel von Inklusion ist der Aspekt der politischen Partizipation bzw. Mitbestimmung im schulischen Kontext zu nennen. Bei der Analyse wurde sichtbar, dass diverse Aspekte von politischer Partizipation explizit aufgeführt werden.

- Schulentwicklung, z. B. „Bei bestimmten Themen werden Schülerinnen und Schüler in die Weiterentwicklung der Schule einbezogen (z. B. Gremien, Mitarbeit in Arbeitsgruppen)“ (LFB_HE);
- Schulleben, z. B. „Mein Kind beteiligt sich an der Gestaltung des Schullebens (z. B. Theateraufführungen, Schulfeste, Konzerte)“ (EFB_HE);
- Schulhausgestaltung, z. B. „Wir Schüler dürfen mitbestimmen, wie unser Schulhaus und unser Schulhof gestaltet werden“ (SFB_BY);

- Schülergremien, z. B. „Die Klassensprecher/innen treffen sich regelmäßig und werden ernst genommen“ (SFB_BE);
- Übernehmen von Verantwortung, z. B. „Übernahme von Verantwortung für die Schule als Ganzes (Streitschlichtung/ Mediation, Schüleraufsicht, Schülerparlament, Initiativen im Gemeinwesen, Lernpatenschaften...)“ (ÜL_ST).

Darüber hinaus gibt es auch allgemein formulierte Items, z. B. „Schülerinnen und Schüler werden nach Verbesserungsvorschlägen gefragt“ (LFB_BY). Generell lässt sich beobachten, dass das Thema politische Partizipation in den Instrumenten aller Bundesländer thematisiert wird. Es wird jedoch nicht thematisiert, dass Kinder und Jugendliche erfahren, dass sie Rechte haben, und, ob sie sich dessen bewusst sind. Dies wäre für Graumann (2012, S. 90) die Operationalisierung der rechtlichen Anerkennungssphäre nach Honneth in inklusiven Settings (nähere Ausführung → Kapitel 2).

Des Weiteren ergibt sich aus der Inklusionsperspektive der Fokus auf die Frage, wer an partizipativen Prozessen beteiligt ist (nähere Ausführung → Kapitel 2). Hierbei zeigt sich, dass nur in wenigen Bundesländern (HH, BE, BW, SH) thematisiert wird, inwiefern alle Schüler_innen die Möglichkeit haben, den schulischen Alltag mitzugestalten, bzw. inwiefern in Bezug auf politische Partizipation die Heterogenität der Schüler_innen berücksichtigt wird. Im Fall von Hamburg steht in den „Merkmale von guter Schule“ (ÜI): „Es gibt eine Kultur, die die Beteiligung aller Schülerinnen und Schüler im Alltag fördert“ (ÜI_HH; siehe auch SFB_BE). So wird hier nicht zwischen verschiedenen Gruppen differenziert, aber es ist ein Bewusstsein vorhanden, dass der Grad an Beteiligung innerhalb einer Lerngruppe aus unterschiedlichen Gründen variieren kann. Des Weiteren wird im Fall von Baden-Württemberg und Schleswig-Holstein die Situation von Schüler_innen mit sonderpädagogischem Förderbedarf explizit benannt:

- „An der Schule gibt es ein abgestimmtes Verfahren wie Schülerinnen und Schüler mit Anspruch auf ein sonderpädagogisches Bildungsangebot dabei unterstützt werden, an der Schulgemeinschaft selbstverständlich teilzunehmen“ (ÜI_BW).
- „Schule ist ein Ort, an dem Formen des sozialen Umgangs, des geregelten Zusammenlebens und der demokratischen Beteiligung erfahren und gelernt werden. Dies gilt auch für die Schülerinnen und Schüler mit sonderpädagogischem Förderbedarf“ (ÜI_SH).

Demnach wird in Schleswig-Holstein und Baden-Württemberg überprüft, inwiefern Schüler_innen mit sonderpädagogischem Förderbedarf politisch partizipieren (können). Hierbei ist jedoch zu berücksichtigen, dass die Gruppe von Schüler_innen mit Förderbedarf als eine relevante Gruppe in den „Merkmale von guter Schule“ aufgeführt wird, aber im Fall von Schleswig-Holstein in den konkreten Instrumenten wie dem Eltern- und Schülerfragebogen nicht mehr zwischen verschiedenen Gruppen differenziert wird. Es bleibt daher unklar, aufgrund welcher Datenlage Aussagen zu diesem Merkmal getroffen werden (im Fall von Baden-Württemberg lagen die Fragebögen nicht vor). Des Weiteren ist zu berücksichtigen, dass auch in diesen beiden Ländern nur die Gruppe von Schüler_innen mit sonderpädagogischem Förderbedarf explizit erwähnt wird und andere Differenzlinien in Bezug auf politische Partizipation nicht explizit genannt werden.

Abgesehen von den oben beschriebenen Fällen findet in den anderen Bundesländern keine Überprüfung dessen statt, inwiefern die Teilhabe aller gewährleistet ist. Jedoch bietet die Befragung von Eltern und Schüler_innen in manchen Bundesländern das Potenzial aussagekräftige Daten in Bezug auf die Teilhabe aller zu erfassen. So wird im Fall von Hessen im Elternfragebogen

überprüft, wie das eigene Kind den schulischen Alltag mitgestalten kann (EFB_HE; siehe auch EFB_BE). Des Weiteren wird im Berliner Schülerfragebogen erfasst, in welcher Weise alle Kinder (SFB_BE/Wahlbereich) bzw. das befragte Kind politisch partizipieren können:

- Alle Kinder: „Bei uns werden alle Kinder ernstgenommen und können aktiv in der Schule mitmachen, egal wie verschieden sie sind“ (SFB_BE/ Wahlbereich).
- Ich-Perspektive des befragten Kindes:
 - „Ich kann mit der Schulleiterin/dem Schulleiter reden, wenn ich das möchte“ (SFB_BE_Pflichtbereich).
 - „Die Schule ermöglicht mir, mich aktiv in das Schulleben einzubringen“ (SFB_BB).

Jedoch besteht in Bezug auf den Eltern- und Schülerfragebogen - wie bereits dargestellt - die Gefahr, dass die Perspektive von einzelnen Schüler_innen verloren geht, wenn bei der Darstellung der Daten nur Durchschnittswerte angegeben werden: So zeigt sich beispielsweise im Kurzbericht des betrachteten Berliner Inspektionsberichts (IB_BE), dass im Qualitätsprofil bei dem Indikator „Beteiligung der Schülerinnen und Schüler und der Eltern“ der Durchschnittswert „A“ (Bewertung bewegt sich zwischen „A“ und „D“) angegeben wurde, aber keine Aufschlüsselung des tatsächlichen Antwortverhaltens gegeben wird.

7.1.3 Ziel von Inklusion: Erfahren von Anerkennung

Als weiteres Ziel von Inklusion ist das Erfahren von intersubjektiver Anerkennung zu nennen. Bei der Analyse kristallisierten sich mehrere Dimensionen heraus, die in diesem Zusammenhang von Relevanz sind:

- Beschreibung, wie die Lehrkraft von Seiten der Schüler_innen behandelt wird, z. B. „Die Schüler gehen wertschätzend mit der Lehrkraft um“ (ÜI_BY).
- Beschreibung, wie die Schüler_innen von Seiten der Lehrkraft behandelt werden, z. B. im Lehrerfragebogen „Ich behandle meine Schülerinnen und Schüler fair“ (LFB_HE).
- Beschreibung des Umgangs zwischen Schüler_innen, z. B. „Die Schülerinnen und Schüler gehen freundlich und rücksichtsvoll miteinander um“ (ÜI_BB) bzw. eine negativ gepolte Abfrage „Manchmal halte ich in der Klasse lieber meinen Mund, weil ich Angst habe, dass sonst wieder jemand lacht“ (SFB_BY).
- Beschreibung davon, wie der Umgang zwischen allen Akteuren innerhalb der Schulgemeinschaft ist, z. B. „Der mitmenschliche Umgang ist in der Schule auf allen Ebenen höflich und wertschätzend“ (ÜI_BE).
- Fehlen von Gewalt/Diskriminierung bzw. Umgang damit, z. B. „Wenn es an meiner Schule Gewalt gibt oder etwas zerstört wird, wird etwas dagegen getan“ (SFB_BE/Wahlbereich) oder „Die Schule reagiert lösungsorientiert auf Konflikte“ (ÜI_BLI).
- Anerkennung der Individualität, z. B. „Alle an Schule Beteiligten werden in ihrer Individualität geachtet“ (ÜI_BY).

In Bezug auf den Themenbereich zeigt sich, dass fast alle Bundesländer auf die gleichen Adjektive zurückgreifen, um diesen zu beschreiben: „respektvoll“, „wertschätzend“, „freundlich“ und „fair“. Dabei tauchen diese Items – bis auf wenige Ausnahmen – in allen untersuchten Instrumenten aller Bundesländer auf, z. B. in den Merkmalen von guter Schule (ÜI_BW; ÜI_BY; ÜI_BE; ÜI_BB; ÜI_HE; ÜI_ST; ÜI_SH) sowie im Schülerfragebogen (SFB_BY; SFB_BE; SFB_BB; SFB_HH; SFB_HE;

SFB_SL; SFB_ST; SFB_SH). Lediglich in drei Fällen gibt es keine Thematisierung im Lehrer- oder Elternfragebogen (LFB_SH; LFB_BB sowie EFB_BB). Zwischen den Bundesländern gibt es geringe Abweichungen, die darin bestehen, welche und wie viele Interaktionen abgefragt werden, z. B. Behandlung der Schüler_innen durch Lehrkraft oder Behandlung der Lehrkraft durch Schüler_innen.

Während in den Fragebögen primär die Einschätzungen der betroffenen Akteure abgefragt werden, wird es im Unterrichtsbeobachtungsbogen den Inspektorinnen und Inspektoren überlassen, den Umgang zwischen schulischen Akteuren aus der Außenseiterperspektive zu evaluieren. Hierbei stellt sich in Bezug auf die Validität die Frage, inwiefern die Aussagen von Schulinspektor_innen, wenn sie im Rahmen von Unterrichtsbeobachtungen erfasst werden, aussagekräftig sind: So ist zum einen zu berücksichtigen, dass die Inspektor_innen sich in allen Bundesländern (außer Bayern) nur für zwanzig Minuten im Unterricht befinden, um die Interaktion zwischen den einzelnen Akteuren zu beobachten. Demnach fehlt es den Inspektor_innen an Vorwissen über das Verhältnis zwischen Schüler_innen und der Lehrkraft. So kann es beispielsweise sein, dass ein bestimmter Umgangston von allen Beteiligten für wertschätzend empfunden wird, aber für die beobachtende Person ungewöhnlich ist. Des Weiteren besteht die Möglichkeit, dass Lehrkräfte in Bezug auf soziale Interaktionen sozial erwünscht handeln und ihr Verhalten vom Alltagsverhalten abweicht. Jedoch könnte in Bezug auf den letzten Punkt argumentiert werden, dass in diesem Fall bei der Auswertung des Unterrichtsbeobachtungsbogens generell nur Bestnoten herauskommen müssten, wenn davon ausgegangen werden würde, dass Lehrkräfte in diesen beobachteten Sequenzen sozial erwünscht handelten. Dies ist nicht der Fall. Des Weiteren sind in Hinblick auf diesen Aspekt die Antwortkategorien im Unterrichtsbeobachtungsbogen zu berücksichtigen. In einigen Bundesländern gibt es zweistufige Antwortkategorien, z. B. „trifft zu/trifft nicht zu“ in Nordrhein-Westfalen, in anderen Bundesländern gibt es vier Niveaustufen, z. B. „1 trifft zu“ / „2 trifft eher zu“ / „3 trifft eher nicht zu“ / „4 trifft nicht zu“ in Sachsen-Anhalt (nähere Ausführung → Kapitel 6).

Eine Ausnahme stellt *Schleswig-Holstein* dar: Hier wird beim Unterrichtsbeobachtungsbogen kein Adjektiv genannt, was sich dadurch erklären lässt, dass der Bogen qualitativ angelegt ist und aus Stichworten besteht.

Darüber hinaus bietet der Schülerfragebogen die Möglichkeit, dass die Perspektive von einzelnen Personen erfasst wird. Jedoch zeigt sich, dass nur wenige Bundesländer dieses Format der Erfassung nutzen, z. B. Differenzierung auf individueller Ebene in Brandenburg: „Meine Lehrerin oder mein Lehrer behandelt mich fair“ (SFB_BB) bzw. „Bei uns werden alle Kinder ernst genommen und können aktiv in der Schule mitmachen, egal wie verschieden sie sind“ (SFB_BE_ Wahlbereich). Jedoch ist auch im Fall der Fragebögen zu berücksichtigen, dass dieser aufgrund des quantitativen Designs und einer begrenzten Antwortskala ähnliche Limitationen aufweist, wie sie eben in Bezug auf den Unterrichtsbeobachtungsbogen dargestellt wurden. Des Weiteren werden die Daten aus den unterschiedlichen Instrumenten bei der Datenaufbereitung – wie bereits oben dargestellt – häufig aggregiert dargestellt. Dagegen wählt *Sachsen-Anhalt* einen anderen Weg: Hier werden Lehrkräfte befragt, inwiefern sie die Qualität pädagogischer Beziehungen als wichtiges Kriterium für Schulqualität ansehen. Dabei geht es also um keine Bestandsaufnahme in Bezug auf pädagogische Beziehungen, sondern um die Frage, inwiefern die Qualität pädagogischer Beziehung eine normative Forderung ist, die von Seiten der Lehrkräfte für relevant erachtet wird.

Bislang ging es um Interaktionen zwischen schulischen Akteuren. Das Anerkennungskonzept von Honneth (1992) geht jedoch einen Schritt weiter im Sinne von Anerkennung von Indi-

vidualität. Dieser Aspekt wird jedoch nur selten explizit thematisiert: „Bei uns werden alle Kinder ernstgenommen und können aktiv in der Schule mitmachen, egal wie verschieden sie sind“ (SFB_BE_Wahlbereich; siehe auch ÜI_BY). Eine alternative Form der Operationalisierung haben in dem Zusammenhang *Bayern* und *Hamburg* gewählt: So werden Lehrkräfte befragt, inwiefern sie der folgenden Aussage zustimmen: „Ich kann an dieser Schule <ich selbst> sein“ (LFB_BY) oder in Hamburg „Die Schulleiterin/der Schulleiter erkennt meine individuellen Bedürfnisse, Fähigkeiten und Ziele“ (LFB_HH). Dies stellt eine Möglichkeit dar zu erfassen, inwiefern Individualität gesehen und anerkannt wird. Dieses Item taucht jedoch nur in Bayern und Hamburg auf und in diesen Fällen auch nur im Lehrerfragebogen und nicht im Schülerfragebogen. Demnach gibt es in keinem der Schülerfragebögen Einschätzungen, inwiefern Schüler_innen den Eindruck haben, in ihrer Individualität wahrgenommen und wertgeschätzt zu werden.

Zusammenfassend lässt sich beobachten, dass bundesländerübergreifend die Interaktionsqualität durch eine Vielzahl von Items erfasst. Dies stellt jedoch nur einen Teilaspekt des Ziels „Erfahren von intersubjektiver Anerkennung“ dar. Um weitere Aspekte, wie Anerkennung als Mensch mit individuellen Rechten, Interessen und Fähigkeiten berücksichtigen zu können, bräuchte es in allen Bundesländern mehr Indikatoren, beispielsweise Selbsteinschätzungsfragen im Schülerfragebogen.

7.1.4 Nennung von Differenzlinien

In den Instrumenten der Schulinspektionen sowie in den „Merkmale von guter Schule“ werden diverse Differenzlinien aufgeführt, die unter folgenden Sammelbegriffen zusammengefasst werden können:

Leistung, z. B. „Meine Lehrerin oder mein Lehrer stellt den leistungsstärkeren Schülerinnen und Schülern schwierigere Aufgaben“ (SFB_BB). Dieser Differenzlinie wird in allen Bundesländern thematisiert und findet sich in fast allen analysierten Instrumenten wieder. Wie bereits in Bezug auf die Operationalisierung von Lernprozessen (→ Abschnitt 7.1.1) angesprochen, gibt es hier unterschiedliche Formen der Differenzierung.

Behinderung/ sonderpädagogischer Förderbedarf: Diese Differenzlinie dient als Sammelbegriff für verschiedene Bezeichnungen (nähere Ausführung → Kapitel 2) und findet sich in mehreren Bundesländern wieder, z. B. „Spezifische Bedürfnisse von Schülerinnen und Schülern mit Anspruch auf ein sonderpädagogisches Bildungsangebot werden gezielt berücksichtigt“ (ÜI_BW; siehe auch ÜI_BY; ÜI_BE; ÜI_ST; ÜI_SH; UBB_NRW; UBB_NI; ÜI_BLI). Hierbei zeigt sich, dass die Bundesländer unterschiedliche Bezeichnungen verwenden:

- Im Fall von *Baden-Württemberg* wird von Schüler_innen mit „sonderpädagogischem Bildungsangebot“ gesprochen bzw. im Fall von *Nordrhein-Westfalen* von sonderpädagogischem Unterstützungsbedarf (UBB_NRW). Dieses bundesländerspezifische Etikett entspricht dem diagnostizierten „sonderpädagogischen Förderbedarf“, welches im Fall von *Sachsen-Anhalt* und *Schleswig-Holstein* verwendet wird, um die Gruppe zu beschreiben. In Bezug auf die BRK impliziert dies jedoch, dass nur eine Teilgruppe der Personen, die unter den Behinderungsbezug der BRK fallen (vgl. Piezunka et al. 2016), in den Instrumenten explizit genannt wird. Beispielsweise fallen darunter nicht diejenigen Personen, die durch die Sozialgesetzbücher geschützt sind.
- Bei der *Bund-Länder-Inspektion* wird von Schüler_innen mit besonderen Bedarfen gesprochen (BLI_ÜI). Dies ist insofern nicht überraschend, da das schulrechtliche Etikett „son-

derpädagogischer Förderbedarf“ in deutschen Schulen im Ausland und somit auch für die Bund-Länder-Inspektion keine praktische Relevanz hat.

- Ansonsten wird in den anderen Bundesländern (*Bayern und Berlin*) explizit von Schüler_innen mit Behinderung und/ oder sonderpädagogischem Förderbedarf gesprochen. Im Fall von *Bayern* wurde zu Beginn der Evaluation von Schüler_innen mit „besonderen physischen und/oder psychischen Voraussetzungen“ (ÜI_BY) gesprochen. Jedoch wurde in 2013 ein Korrekturblatt veröffentlicht. Danach wurde von Schüler_innen „mit Behinderungen bzw. sonderpädagogischem Förderbedarf“ gesprochen.¹
- In *Niedersachsen* zeigt sich im Unterrichtsbeobachtungsbogen, dass Schüler_innen mit Behinderung nicht als homogene Gruppe gesehen werden. Stattdessen wird in Abhängigkeit von dem jeweiligen Indikator nochmal stärker differenziert, z. B. wenn es um die Sprachkompetenz geht, gibt es eine ausführliche Erläuterung zum Indikator: „Inklusion: „Der Unterricht berücksichtigt Schülerinnen und Schüler, die nur eingeschränkt sprechen oder Sprache verstehen können“ (UBB_NI). Wenn es dagegen um eigenverantwortliche Arbeitsprozesse geht: „Inklusion (...): Das Maß der erforderlichen Unterstützung ist abhängig vom Grad der Beeinträchtigung“ (UBB_NI). Demnach wird in den Instrumenten deutlich gemacht, dass sich aus dem Attribut „diagnostizierte Behinderung“ unterschiedliche Bedarfe ergeben können.
- Im Fall von Brandenburg, Hessen und Saarland werden „Behinderung“ bzw. sonderpädagogischer Förderbedarf in den verfügbaren Instrumenten sowie in den „Merkmale von guter Schule“ nicht genannt und im Fall von Hamburg wird es in den Anmerkungen zum UBB, aber nicht in den konkreten Instrumenten des UBB genannt (UBB_HH).

Es gibt sehr unterschiedliche Formate, wie „Behinderung“ bzw. „sonderpädagogischer Förderbedarf“ als Differenzlinie in den Instrumenten thematisiert wird: In den meisten Fällen werden Schüler_innen mit Behinderung bzw. sonderpädagogischem Förderbedarf als Subjekte gesehen und es geht um die Frage, ob ihre gleichberechtigte Teilhabe gewährleistet ist (z. B. LFB_BE) oder inwiefern sie die Unterstützung erhalten, die sie benötigen (ÜI_BW; ÜI_BE; ÜI_SH; ÜI_ST; UBB_NRW; UBB_NI).

Bayern weicht insofern von den anderen Bundesländern ab, als im Lehrerfragebogen folgende Einschätzung abgefragt wird: „Ich habe den Eindruck, dass die Schüler meiner Klasse sich in die Situation von Schülern mit Behinderungen bzw. mit sonderpädagogischem Förderbedarf hineinversetzen können“ (LFB_BY). Durch diese Frage wird der Gruppe von Schüler_innen mit Behinderung im Lehrerfragebogen eine Sonderrolle zugeschrieben. Das „Hineinversetzen“ kann als Abfrage davon gelesen werden, inwiefern die als nicht behindert diagnostizierten Schüler_innen sich mit der Situation dieser Gruppe auseinandergesetzt haben und in der Lage sind, die Interessen und Bedürfnisse der jeweiligen Schüler_innen in ihrem eigenen Handeln zu berücksichtigen. Zugleich impliziert dieser Indikator auch, dass die Gruppe von Schüler_innen mit diagnostizierter Behinderung als von der Norm abweichende Gruppe konstruiert wird. Diese Wirkung wird dadurch verstärkt, dass es die einzige Gruppe ist, die auf diese Art und Weise explizit genannt wird. Dabei lässt die Aussage „Hineinversetzen in die Situation“ darüber hinaus auch Raum für stigmatisierende Zuschreibungen, die auf defizitorientierenden Haltungen basieren. Dies wäre der Fall, wenn „Hineinversetzen in die Situation“ damit einhergeht, dass Schüler_innen mit diagnostizierter Behinderung als Gruppe gesehen wird, die möglichst geschont werden sollte oder der ihr Recht auf Selbstbestimmung abgesprochen wird. Weitere Aussagen in Hinblick auf Schüler_innen mit Behinderung gibt es im

¹ Es waren für die Autorin dieser Arbeit keine Informationen zugänglich, wodurch die Umbenennung initiiert wurde.

bayerischen Schülerfragebogen: „Wir sprechen im Unterricht über den Alltag von Schülern mit Behinderung“ (SFB_BY) sowie „Ich kann mir gut vorstellen, dass behinderte und nichtbehinderte Schüler gemeinsam unterrichtet werden“ (SFB_BY). Indem der Einstellungsaspekt im Rahmen des Schülerfragebogens abgefragt wird, wird den Schüler_innen signalisiert, dass sie sich auch kritisch dazu positionieren können, dass Schüler_innen mit und ohne Behinderung zusammen unterrichtet werden. Eine Positionierung gegen den gemeinsamen Unterricht von Schüler_innen mit und ohne zugeschriebene Behinderung erscheint somit als legitime Antwortkategorie. Hierbei ist zu berücksichtigen, dass die kritische Positionierung zum gemeinsamen Unterricht im öffentlichen Diskurs keine Seltenheit darstellt: Beispielsweise in einer Studie, die vom Verband Bildung und Erziehung e.V. in Auftrag gegeben wurde, sind 42 Prozent der Lehrkräfte der Meinung, dass Kinder mit diagnostizierter Behinderung an einer Förderschule besser aufgehoben wären (z. B. Verband Bildung und Erziehung e.V. 2017). Demnach könnte auch argumentiert werden, dass diese Abfrage der bayerischen Schulinspektion es ermöglicht, Informationen darüber zu erhalten, welche Einstellungen die Schüler_innen gegenüber Inklusion haben. Ausgehend von der Frage, wie Inklusion – wie es im Rahmen dieser Arbeit verstanden wird (→ Abschnitt 2.2) – in den Instrumenten berücksichtigt wird, wäre zu argumentieren, dass dieses Item keine eindeutige Positionierung von Seiten der Schulinspektion dazu ist, dass der gemeinsame Unterricht von Schüler_innen mit und ohne diagnostizierte Behinderung ein Qualitätsmerkmal von Schule darstellt. Des Weiteren spielen bei der Beantwortung der Frage die Mehrheitsverhältnisse innerhalb der Schülerschaft eine Rolle: So sind Schüler_innen ohne Behinderung im Vergleich zu Schüler_innen mit Behinderung an Regelschulen zahlenmäßig stärker vertreten, das heißt die Entscheidungsmacht darüber, ob gemeinsamer Unterricht zu befürworten ist, wird den Schüler_innen ohne Behinderung überlassen. Die oben beschriebenen Indikatoren aus Bayern sind aus Inklusionsperspektive insofern kritisch zu hinterfragen, da dadurch von Seiten der Schulinspektion die Gruppe von Schüler_innen mit Behinderung als von der Norm abweichend konstruiert wird. Des Weiteren ist zu berücksichtigen, dass es im Fall von Bayern neben den oben beschriebenen Aussagen keine zusätzlichen Indikatoren gibt, die Schüler_innen mit diagnostizierter Behinderung als Subjekte mit bestimmten Rechten verstehen und somit nicht überprüft wird, inwiefern deren Rechte geschützt sind.

Sprache/ Wahrnehmung von Anderen als Nicht-Deutsche: In diesem Fall werden in der empirischen Analyse unter dem oben aufgeführten Sammelbegriff verschiedene Konzepte zusammengefasst, die im Alltag häufig synonym verwendet werden, obwohl sich dahinter unterschiedliche Gruppen verbergen. Eine explizite Nennung findet sich in vier Bundesländern: Baden-Württemberg, Berlin, Schleswig-Holstein und Hamburg.

- Im Fall von *Baden-Württemberg* ist ein Anspruch, dass die sozialen und kulturellen Voraussetzungen der Schülerschaft berücksichtigt werden (ÜI_BW).
- In *Berlin* liegt der Fokus auf der Förderung von Schüler_innen im sprachlichen Bereich, z. B. „Kindern, die Probleme mit der Sprache haben, wird bei uns besonders geholfen“ (SFB_BE; LFB_BE; EFB_BE).
- In *Schleswig-Holstein* findet sich in den Erläuterungen die Abfrage, inwiefern Schüler_innen unabhängig von ihrem Migrationshintergrund an der Schule aufgenommen werden (ÜI_SH).
- Im Fall von *Hamburg* wird im Elternfragebogen abgefragt, inwiefern auf die „kulturellen Unterschiede der Schülerinnen und Schüler eingegangen“ (EFB_HH) wird und inwiefern auf die „sprachlichen Unterschiede der Schülerinnen und Schüler eingegangen“ (EFB_HH; siehe auch Anmerkungen zum UBB_HH) wird. Aus Inklusionsperspektive ist positiv zu bewerten, dass hinsichtlich kultureller und sprachlicher Unterschiede nicht von einer dominierenden

Kultur ausgegangen wird und alle anderen kulturellen Zugehörigkeiten als Abweichungen von der Norm gesehen bzw. defizitäre Zuschreibungen gemacht werden. In diesem Zusammenhang stellt sich jedoch die Frage, was mit „kulturellen Unterschieden“ gemeint ist. Dahinter könnte sich beispielsweise verbergen, dass manche Schüler_innen dadurch geprägt sind, dass Angehörige ihrer Familie eine Migrationsgeschichte haben. Des Weiteren können darunter auch unterschiedliche Religionszugehörigkeiten verstanden werden. Es fehlt daher an Angaben, welche konkreten kulturellen Unterschiede im schulischen Kontext von Seiten der Schulinspektion für relevant erachtet werden. Darüber hinaus fehlt es bei der Formulierung „Eingehen auf Unterschiede“ an Informationen darüber, wie schulische Akteure auf kulturelle Unterschiede eingehen können bzw. sollen. Beispielsweise kann das Eingehen auf kulturelle Unterschiede implizieren, dass im Rahmen eines interkulturellen Frühstücks stereotype Zuschreibungen in Bezug auf bestimmte Herkunftsländer reproduziert werden. Die Berücksichtigung von kulturellen Unterschieden kann jedoch auch implizieren, dass Lehrkräfte über Wissen zu bestimmten Religionen verfügen, z. B. welche Feiertage für die Schüler_innen von Relevanz sein könnten und welche Traditionen damit einhergehen. Wie bereits oben dargestellt, besteht auch hier bei der Beantwortung der Frage die Gefahr, dass die Einschätzungen von zahlenmäßig kleinen Gruppen nicht wahrgenommen werden, weil sie an vielen Schulen zahlenmäßig unterlegen sind. Um dies zu illustrieren: Eltern mit Deutsch als Muttersprache haben möglicherweise den Eindruck, dass an ihrer Schule ausreichend auf die sprachlichen Unterschiede eingegangen wird. Es kann jedoch sein, dass die Einschätzungen von Eltern von Schüler_innen mit Deutsch als Zweitsprache davon abweichen. Ihre Perspektive würde jedoch kaum gesehen werden, da sie innerhalb einer Klasse meistens eine kleine Gruppe von Personen repräsentieren bzw. nicht identifiziert werden kann, wer welche Antworten gegeben hat.

- Neben diesen drei Fällen ist noch die *Bund-Länder-Inspektion* zu nennen: So gibt es in den „Merkmale von guter Schule“ Indikatoren, die auf „Deutsch als Fremdsprache“ sowie auf die Differenzlinie „Kultur“ anspielen: „Der Unterricht fördert gezielt die Entwicklung der Deutschkompetenz“ (ÜI_BLI) sowie „Die schulische Arbeit erzielt die Herausbildung interkultureller Kompetenz“ (ÜI_BLI). Im Gegensatz zu den oben beschriebenen Fällen geht es hier nicht um die Situation von einzelnen Schülergruppen, sondern um den Wert von „interkultureller Kompetenz“ und den Erwerb von Deutsch als Fremdsprache für die Mehrheit der Schüler_innen. So werden an Auslandsschulen 10.000 Schüler_innen mit deutscher Staatsbürgerschaft und 57.500 nicht-deutsche Schüler_innen unterrichtet (Ministerium für Kultus, Jugend und Sport Baden-Württemberg 2017). Der Fokus auf interkulturelle Kompetenz ergibt sich daher aus der Zielsetzung von Auslandsschulen: So werden Auslandsschulen als „Träger der Auswärtigen Kulturpolitik“ (ebd.) gesehen und ihre Arbeit zielt darauf ab, dass möglichst viele Schüler_innen nach dem Schulbesuch weiterhin Bezüge zu Deutschland aufrechterhalten, z. B. durch die Aufnahme eines Studiums in Deutschland. Der Erwerb von interkultureller Kompetenz würde sich demnach darin widerspiegeln, Kenntnisse in Bezug auf Deutschland zu vermitteln und Bezüge zum jeweiligen Gastgeberland herzustellen.

Zusammenfassend zeigt sich, dass nur wenige Bundesländer sprachliche Unterschiede bzw. die kulturelle Verortung als Differenzlinien explizit benennen. Des Weiteren zeigt sich bundesländerübergreifend, dass Diskriminierung, die stattfindet, weil eine Person von anderen als Nicht-Deutsch wahrgenommen wird, in den Instrumenten aller Bundesländer nicht thematisiert wird. Dies ist insofern überraschend, als es eine Vielzahl von empirischen Studien gibt,

wonach zumindest die Gruppe von Schüler_innen mit Migrationshintergrund² im schulischen Kontext diskriminiert wird, z. B. durch negative Zuschreibungen (z. B. Radtke und Gomolla 2009) oder in der Leistungsbewertung (vgl. Sprietsma 2009).

Soziale Herkunft: In Bezug auf soziale Herkunft zeigt sich ein ähnliches Phänomen wie bei der vorherigen Differenzlinie. Diese wird nur in Baden-Württemberg, Berlin, Hamburg, Schleswig-Holstein bzw. in einem anderen Format auch in Sachsen-Anhalt explizit erwähnt.

- Dabei geht es in *Berlin* um die Frage, inwiefern die gleichberechtigte Teilhabe unabhängig von Herkunft (LFB_BE; EFB_BE) gewährleistet ist. Im Fall von *Schleswig-Holstein* wird erfasst, ob die Schule unabhängig von sozialer Herkunft alle Schüler_innen aufnimmt und zu erfolgreichen Abschlüssen führt (ÜI_SH).
- Im Fall von *Hamburg* sind es die Eltern, die befragt werden, inwiefern auf die sozialen Unterschiede der Schüler_innen eingegangen wird (EFB_HH; siehe auch Anmerkungen zum UBB_HH). Hier zeigen sich bei der Operationalisierung und Darstellung der Ergebnisse aufgrund der abstrakten Formulierung des Indikators ähnliche Herausforderungen wie bei der vorherigen Differenzlinie (→ Berücksichtigung von „sprachlichen Unterschieden“).
- In *Baden-Württemberg* geht es darum, ob in den Unterrichtsangeboten die sozialen Voraussetzungen der Schüler_innen berücksichtigt werden (ÜI_BW).
- In *Sachsen-Anhalt* werden Lehrkräfte befragt, inwiefern das Lernen durch bestimmte Aspekte beeinträchtigt wird. Eine mögliche Antwortkategorie ist: „Schülerinnen und Schüler aus sozial benachteiligten Verhältnissen“ (LFB_ST). Während es in Baden-Württemberg, Berlin und Hamburg also die Pädagog_innen sind, die in der Verantwortung gesehen werden, die Teilhabe unabhängig von der sozialen Herkunft zu gewährleisten, werden Schüler_innen aus sozial benachteiligten Verhältnissen im Fall von Sachsen-Anhalt als mögliche Erklärung für auftretende Probleme und Beeinträchtigungen gesehen.
- In anderen Bundesländern spielt „soziale Herkunft“ als Differenzlinie keine Rolle in den Instrumenten.

Gender: Gender wird in vier Bundesländern als Differenzlinie explizit benannt: Niedersachsen, Nordrhein-Westfalen, Schleswig-Holstein und Berlin. Dabei werden in jedem Bundesland andere inhaltliche Aspekte angesprochen und des Weiteren unterscheidet es sich auch, welche Relevanz dem Thema beigemessen wird.

- In *Niedersachsen* wird als ein mögliches Beispiel für den Indikator „Kooperationen eingehen“ aufgeführt, inwiefern es Kooperationsangebote zum Thema Gender Mainstreaming gibt (ÜI_NI). Demnach ist es kein eigener Indikator, sondern dient nur als Illustration für den Themenbereich Kooperationen. Ein ähnlicher Fall zeigt sich in Schleswig-Holstein. Dort findet sich in den Erläuterungen zu den Merkmalen von guter Schule: „Die Schule nimmt Schülerinnen und Schüler unabhängig von Geschlecht, sozialer Herkunft, Migrationshintergrund oder Behinderung auf und führt sie zu erfolgreichen Abschlüssen“ (ÜI_SH).
- Im Fall von *Nordrhein-Westfalen* gibt es ein Kriterium, inwiefern die Schule ein Konzept zur Gender-Mainstreaming-Erziehung verwirklicht (ÜI_NRW_Wahlbereich). Hierbei ist jedoch zu berücksichtigen, dass dies im Wahlbereich aufgeführt wird, das heißt die Schulen können

² In bisherigen empirischen Studien wird überwiegend auf die Definition des Statistischen Bundesamts zurückgegriffen und es fehlt bislang an Studien zu der Gruppe von Schüler_innen, die von anderen als „Nicht-Deutsche“ wahrgenommen wird. Der Verweis auf die Klassifizierung „Migrationshintergrund“ wird im Rahmen dieser Arbeit gemacht, weil zu erwarten ist, dass es hier große Überschneidungen gibt.

entscheiden, ob sie in dem Bereich evaluiert werden möchten. Des Weiteren wird im Unterrichtsbeobachtungsbogen Folgendes erfasst: „Die Lehrkraft bezieht Mädchen und Jungen gleichermaßen in den Unterricht ein“ (UBB_NRW).

- Im Fall von *Berlin* werden Lehrkräfte befragt, inwiefern in Bezug auf Geschlecht die gleichberechtigte Teilhabe aller gewährleistet ist (LFB_BE_Indikator Inklusion).

Zusammenfassend lässt sich beobachten, dass Gender als Differenzlinie in vielen Bundesländern nicht explizit genannt wird. In den oben beschriebenen Fällen ist zu berücksichtigen, dass Gender als Differenzlinie zum Teil nur auf der Beispielebene genannt wird bzw. ein Wahlkriterium darstellt. Des Weiteren zeigt sich bei der Operationalisierung, dass im Fall von Nordrhein-Westfalen auf die dichotome Unterscheidung zwischen männlich/weiblich zurückgegriffen wird, d. h. andere Formen der sexuellen Identität, z. B. intersexuell, werden nicht explizit genannt. Darüber hinaus wird auch nicht zwischen Fremd- und Selbstzuschreibung unterschieden: So kann es durchaus sein, dass Fremd- und Selbstzuschreibungen in Bezug auf die sexuelle Identität voneinander abweichen.

Private Ausnahmesituation: Ein weiteres Attribut von Schüler_innen, welches in den Instrumenten vieler Bundesländer aufgeführt wird, sind „private Ausnahmesituation“ oder auftretende Probleme. Dies taucht in den „Merkmale von guter Schule“ auf (ÜI_BY; ÜI_BE; ÜI_HE; ÜI_NRW; ÜI_NI; ÜI_ST; ÜI_BLI), im Elternfragebogen (EFB_BY; EFB_BE; EFB_ST), im Lehrerfragebogen (LFB_BY; LFB_BE; LFB_HE; LFB_ST) sowie im Schülerfragebogen (SFB_BE; SFB_HE; SFB_SL). Dabei drehen sich alle Indikatoren um die Frage, inwiefern die Schüler_innen z. B. durch bestehende Beratungsangebote oder durch einzelne Lehrkräfte Unterstützung erhalten, wenn sie sich in Problemsituationen befinden.

Abschließend können in Bezug auf die Berücksichtigung von Differenzlinien folgende Aspekte festgestellt werden: Bundesländerübergreifend lässt sich als Muster beobachten, dass im Unterrichtsbeobachtungsbogen (bis auf Geschlecht in NRW) nur „Leistung“ als relevante Differenzlinie aufgeführt wird. Die Relevanz der Differenzlinie „Leistung“ zeigt sich darüber hinaus darin, dass dieser Aspekt in allen Bundesländern explizit benannt wird. Dies ist in Hinblick auf die gesellschaftlichen Funktionen von Schule wenig überraschend. Es gibt jedoch Unterschiede darin, wie Leistung operationalisiert wird. So greifen viele Bundesländer auf dichotome Unterscheidungen zurück, z. B. leistungsstark vs. leistungsschwach, während andere Bundesländer in manchen Indikatoren deutlich machen, dass es ein Spektrum an Leistungsfähigkeit gibt. Des Weiteren zeigt sich, dass andere Differenzlinien, wie soziale Herkunft oder Behinderung, nur in wenigen Bundesländern explizit aufgeführt werden. Bei der Operationalisierung der Differenzlinien werden zum Teil recht abstrakte Beschreibungen verwendet, z. B. „kulturelle Unterschiede“ (EFB_HH). „Behinderung“ taucht in mehreren Bundesländern als Differenzlinie auf. Jedoch wird nur im Fall von *Niedersachsen* explizit thematisiert, dass diese Gruppe von Schüler_innen in sich sehr heterogen ist (UBB_NI).

Weitere Differenzlinien, die bezüglich Teilhabe und Diskriminierung von Relevanz sind, wie sexuelle Orientierung oder Fremdwahrnehmung als Nicht-Deutsch, werden in den Instrumenten der Schulinspektion gar nicht benannt. Dabei fehlen insbesondere Differenzlinien, die in Kapitel 2 unter das dritte Kriterium fallen: Demnach sind Attribute von Relevanz, die von vielen als von der Norm abweichend angesehen werden. Bei der expliziten Verwendung von Differenzlinien ist auch die Darstellung der Daten zu berücksichtigen: So besteht bei vielen Formen der Operationalisierung die Gefahr, dass die Perspektive von zahlenmäßig kleinen Gruppen, z. B. Schüler_innen mit sonderpädagogischem Förderbedarf, bei der Darstellung von Durchschnittswerten nicht ausreichend berücksichtigt wird. Ein weiterer Aspekt, der im

Zusammenhang mit Differenzlinien in den Instrumenten einbezogen wird, ist die Unterscheidung zwischen Fremd- und Selbstzuschreibung (nähere Ausführung → Abschnitt 2.2). So kann es beispielsweise sein, dass diese voneinander abweichen bzw. dass die Fremdzuschreibung, z. B. Schüler_in mit Migrationshintergrund, kein identitätsstiftendes Merkmal für die jeweilige Person ist und daher auch nicht der Wunsch besteht, dass dieser Aspekt im Schulleben eine große Rolle spielt. Der Aspekt von nicht erwünschten Fremdzuschreibungen könnte im Rahmen des Schülerfragebogens erfasst werden, indem Schüler_innen befragt werden, ob ihnen stereotype Zuschreibungen im schulischen Alltag begegnen. Hiermit würde ein Kritikpunkt, der im Zusammenhang mit dem Honneth'schen Anerkennungskonzepts genannt wird, aufgegriffen werden (nähere Ausführung → Kapitel 2), indem Wahrnehmung und Wertschätzung von Individualität nicht per se als etwas Positives verstanden werden, sondern Raum dafür da ist, kritisch zu hinterfragen, inwiefern durch positiv intendierte Wahrnehmung von Individualität stereotype Zuschreibungen und somit die Reproduktion von Machtverhältnissen einhergehen.

Darüber hinaus findet sich in den betrachteten Instrumenten auch nicht die Berücksichtigung einer intersektionalen Perspektive wieder. So könnte argumentiert werden, dass der Fokus auf individuelle Differenzierung im Lernprozess diesen Aspekt beinhaltet. Jedoch spielt der Fokus auf das Individuum in den Instrumenten nur eine geringe Rolle und der Aspekt der Intersektionalität wird in den Instrumenten nicht explizit gemacht. Dies ist insofern von Relevanz, als bestimmte Bedürfnisse von Schüler_innen bzw. auch die Außenwahrnehmung häufig in Abhängigkeit von der Verschränkung von Differenzlinien variieren. Beispielsweise konnte unter anderem Ispa-Landa nachweisen, dass die Außenwahrnehmung in Bezug auf die Verschränkung von „race“ und Gender in US-amerikanischen Schulen variiert (vgl. Ispa-Landa 2013).

7.1.5 Berücksichtigung von Inklusion als normative Forderung

Im Rahmen der Analyse zeigte sich, dass sich auf Bundeslandebene keine systematischen Muster herausbilden lassen, die über die einzelnen Aspekte von Inklusion hinausgehen. Demnach bietet sich in diesem Fall eine Typenbildung nicht an. So lassen sich lediglich hinsichtlich einzelner Aspekte unterschiedliche Typen herausbilden.

- Bei der „Differenzierung im Lernprozess“ zeigt sich als Muster, dass dieser Themenbereich in fast allen Bundesländern thematisiert bzw. operationalisiert wird, aber es fehlt häufig an genauen Angaben darüber, wie die Lerngruppe unterteilt bzw. deren Heterogenität wahrgenommen werden soll. So wird Differenzierung auf Individualebene selten explizit in den Instrumenten genannt. Das Potenzial Selbsteinschätzungen auf Individualebene durch den Schülerfragebogen zu erfassen wird nur im Fall von *Schleswig-Holstein* genutzt, indem in der Ergebnisdarstellung neben dem Durchschnittswert die Verteilung auf die jeweiligen Antwortkategorien aufgeführt wird. Es fehlt jedoch auch hier an Angaben darüber, ob es hier systematische Unterschiede auf der Grundlage von bestimmten Differenzlinien gibt.
- Bei dem Ziel „Erfahren von intersubjektiver Anerkennung“ ergibt sich, dass die Interaktionsebene in allen Bundesländern thematisiert wird, z. B. Umgang zwischen Schüler_innen und Lehrkraft. Mit Verweis auf Honneth ist eine Sphäre das Anerkennen von individuellen Merkmalen und Fähigkeiten. Dies könnte beispielsweise im Schülerfragebogen erhoben werden, indem abgefragt wird, ob Schüler_innen in ihrer Individualität gesehen werden und wertgeschätzt fühlen. Einen möglichen Indikator findet sich im Fall von Berlin, wobei hier

die Schüler_innen nicht Einschätzungen über sich, sondern über die gesamte Lerngruppe („alle Kinder“) treffen sollen.

- In Bezug auf das Ziel „politische Partizipation aller“ zeigt sich bei der Analyse, dass die Bundesländer unterschiedliche Aspekte von politischer Partizipation erfassen. Jedoch gibt es nur zwei Bundesländer, Schleswig-Holstein und Baden-Württemberg, in denen explizit (im Fall von Hamburg und Berlin implizit) abgefragt wird, inwiefern auch die Partizipation von Teilgruppen, z. B. Schüler_innen mit sonderpädagogischem Förderbedarf, sichergestellt wird. Hierbei ist nicht ersichtlich, warum Schüler_innen mit sonderpädagogischem Förderbedarf als Gruppe explizit herausgegriffen werden.
- Bezüglich der Nennung von einzelnen Differenzlinien zeigt sich als Muster, dass neben den Differenzlinien „Leistungsfähigkeit“ und „Behinderung“ in den meisten Bundesländern keine anderen Differenzlinien genannt werden, die z. B. bei Diskriminierung von Relevanz sind. Eine Ausnahme stellen Berlin, Baden-Württemberg, und Hamburg (sowie Schleswig-Holstein in den Erläuterungen) dar, wo die Berücksichtigung von kulturellen/sprachlichen und sozialen Unterschieden erfasst wird.

Einige Bundesländer stechen in einzelnen Aspekten heraus: *Baden-Württemberg*, *Berlin* und *Hamburg* führen eine Vielzahl von unterschiedlichen Differenzlinien auf, die in Bezug auf verschiedene Ziele von Inklusion wie politische Partizipation oder Differenzierung im Unterricht von Relevanz sind. Des Weiteren entsteht im Fall von *Brandenburg* und *Saarland* der Eindruck, dass Differenzierung sowie die explizite Berücksichtigung von bestimmten Gruppen (auch keine Nennung von Behinderung) im Vergleich zu den anderen Bundesländern eine geringere Rolle spielen. *Schleswig-Holstein* stellt eine Ausnahme bezüglich seiner Instrumente dar, als dass es einen qualitativen Unterrichtsbeobachtungsbogen verwendet. Demnach können komplexe Konzepte beschrieben werden. Ein Unterrichtsbeobachtungsbogen, der auf Stichworten basiert, hat jedoch zur Folge, dass anzustrebende Normen nur zum Teil explizit gemacht werden können und den jeweiligen Inspektor_innen wird mehr Autonomie im Bewertungsprozess zugeschrieben.

7.2 Verwendung des Inklusionsbegriffs

Im Folgenden wird nun dargestellt, wie der Inklusionsbegriff explizit in den Instrumenten der Schulinspektionen verwendet wird. Hierbei wurde zusätzlich noch der Qualitätsrahmen berücksichtigt, da es eher unüblich ist, in Messinstrumenten Definitionen aufzuführen.

7.2.1 Verständnis von Inklusion

Im ersten Schritt wurde näher betrachtet, ob die Schulinspektionen in ihren Instrumenten und weiteren Unterlagen Definitionen von Inklusion nennen und wie sich diese voneinander unterscheiden (für eine Übersicht → Tab. 7.2). So wurden im Fall von vier Bundesländern Definitionen in den Messinstrumenten gefunden: Bayern, Baden-Württemberg, Berlin und Hamburg.

Im Fall von *Hamburg* wird im Eltern- und Lehrerfragebogen auf die Definition der Salamanca-Konferenz verwiesen und diese wird explizit aufgeführt:

„Laut Erklärung der UNESCO-Konferenz von Salamanca im Jahr 1994 umfasst Inklusion alle Kinder, unabhängig von ihren physischen, intellektuellen, sozialen, emotionalen, sprachlichen oder anderen

Fähigkeiten und schließt behinderte und begabte Kinder, Kinder von sprachlichen, kulturellen oder ethnischen Minoritäten sowie Kinder von anders benachteiligten Randgruppen oder -gebieten ein“ (HH_LFB; siehe auch HH_EFB; Behörde für Schule und Berufsbildung Hamburg 2012).

Demnach positioniert sich die Hamburger Schulinspektion deutlich zu einem weiten Inklusionsbegriff. Es fehlen jedoch nähere Angaben darüber, welche Ziele mit Inklusion verfolgt werden.

Im Fall von *Berlin* wird im Handbuch der Inspektion (Berliner Senatsverwaltung für Bildung, Jugend und Wissenschaft 2012) sowie im Handlungsrahmen (2012) auf einen weiten Inklusionsbegriff zurückgegriffen. Mit Verweis auf die BRK wird näher dargestellt, welche Ziele mit Inklusion im Allgemeinen bzw. im Bildungsbereich verknüpft werden:

„Die Staaten, die die BRK unterzeichnet haben, leisten dies, indem sie Ansprüche auf Selbstbestimmung, Diskriminierungsfreiheit und gleichberechtigte gesellschaftliche Teilhabe formulieren, sie rechtsverbindlich verankern und mit möglichst wirksamen Durchsetzungsformen verknüpfen. Für den Bildungsbereich bedeutet das, allen Kindern und Jugendlichen uneingeschränkt Zugang zur allgemeinen Schule zu ermöglichen. Ziele sind eine umfassende Persönlichkeitsentwicklung, der Erwerb eines Schulabschlusses entsprechend den individuellen Möglichkeiten und die gleichberechtigte und chancengleiche Teilhabe am Leben in der Gesellschaft“ (Berliner Senatsverwaltung für Bildung, Jugend und Wissenschaft 2012, S. 52).

Demnach zeigen sich hier eindeutig Anknüpfungspunkte zu dem Inklusionsverständnis, welches in dieser Arbeit verwendet wird (→ Abschnitt 2.2), da sich Inklusion auf alle Schüler_innen bezieht und als Ziel von Inklusion unter anderem Diskriminierungsfreiheit genannt wird. Des Weiteren wird im Fall von Berlin explizit auf ein soziales Modell von Behinderung verwiesen:

„Die Behindertenrechtskonvention markiert einen grundlegenden Wechsel, indem sie den traditionellen, primär an Defiziten der Betroffenen orientierten Ansatz durch einen <diversity-Ansatz> (Anderssein und Verschiedenheit wertschätzen) ersetzt. Sie signalisiert eine Abkehr von einer Behindertenpolitik, die primär auf Fürsorge und Ausgleich von Defiziten abzielt“ (ebd., S. 52).

Im Fall von *Baden-Württemberg* werden im „Qualitätsrahmen zur Fremdevaluation“ (Landesinstitut für Schulentwicklung Baden-Württemberg 2015) zwei Definitionen von Inklusion aufgeführt:

„Inhaltlich wurde das Thema Heterogenität von Schülerinnen und Schülern und **Inklusion im weiteren Sinne** quer über den gesamten Qualitätsrahmen in entsprechenden Merkmalen auf der Ebene der Qualitätsstandards bzw. Indikatoren eingearbeitet. Für die **Inklusion im engeren Sinne** (Schülerinnen und Schüler mit Anspruch auf ein sonderpädagogisches Bildungsangebot) gibt es in entsprechenden Merkmalen zusätzliche Indikatoren (...)“ (ebd., S. 5).

Auch hier liegt bei der Definition von Inklusion der Fokus auf der Frage, welche Zielgruppe darunter fällt. Es wird zwar explizit formuliert, dass unter Inklusion der Umgang mit heterogenen Lerngruppen und somit alle Schüler_innen als Zielgruppe verstanden werden. Im Rahmen der Instrumente wird jedoch auf das zweite Inklusionsverständnis („im engeren Sinne“) zurückgegriffen. Aus normgebender Perspektive kann dieser Textbaustein auch als Signal gelesen werden, dass hier von Seiten der Schulinspektion in Bezug auf die Zielgruppe von Inklusion gegenüber den Schulen keine Vorgaben gemacht werden.

In *Bayern* liegt der Fokus auf der Teilhabe von Schüler_innen mit diagnostizierter Behinderung: „Beim Kriterium Förderung der Integration/ Inklusion geht es um die Frage, inwieweit

die Schule auf die Teilhabe von Schülern mit besonderen physischen und/oder psychischen Voraussetzungen vorbereitet ist“ (Bayerisches Staatsministerium für Unterricht und Kultus 2010, S. 18). Als Ziel von Inklusion wird auf das Konzept der „Teilhabe“ verwiesen, welches im Rahmen der Definition nicht näher ausgeführt wird.

Insgesamt wurde in den Instrumenten von vier Bundesländern eine explizite Definition von Inklusion gefunden. Die oben beschriebenen Definitionen unterscheiden sich darin, wer als Zielgruppe von Inklusion benannt wird. So lässt sich grob zwischen einem engen und einem weiten Inklusionsbegriff unterscheiden, wobei in Baden-Württemberg im Handbuch keine explizite Festlegung getroffen wird, welche Definition für die Schulinspektion gilt. Berlin sticht insofern heraus, als die Ziele von Inklusion in Bezug auf den Bildungsbereich konkretisiert werden und explizit auf ein soziales Modell von Behinderung verwiesen wird. Weitere Hinweise darauf, was mit Inklusion gemeint ist, lassen sich sowohl in den oben beschriebenen vier Fällen als auch in den anderen Bundesländern daraus ablesen, in welchem Zusammenhang Inklusion als Begriff in den Instrumenten verwendet wird.

In vier weiteren Fällen (BLI, NI, NRW, SH) wird Inklusion als Begriff zwar verwendet, es wird aber in den analysierten Materialien der Schulinspektionen nicht definiert, was unter Inklusion verstanden wird. Im Fall von *Niedersachsen* lassen sich aus dem Kontext der Nennung Hinweise ablesen: So wird Inklusion als Begriff im Zusammenhang mit Förderzentren bzw. mit sonderpädagogischem Förderbedarf genannt (Niedersächsisches Landesinstitut für schulische Qualitätsentwicklung 2015, S. 37). Dies gibt Hinweise darauf, dass unter Inklusion der gemeinsame Unterricht von Schüler_innen mit und ohne Behinderung verstanden wird. In Ergänzung wird im Orientierungsrahmen folgendermaßen von Inklusion gesprochen: „die Entwicklung zur inklusiven Schule, in der für alle Schülerinnen und Schüler des regionalen Umfelds ein barrierefreier und gleichberechtigter Zugang eröffnet wird (...)“ (Niedersächsisches Kultusministerium 2014, S. 12). Ein ähnliches Verständnis liegt bei der *Bund-Länder-Inspektion* nahe, da drei von fünf Indikatoren unter dem Qualitätsbereich „inklusive arbeiten“ auf „Schüler mit besonderen Bedarfen“ bzw. Barrierefreiheit (ÜI_BLI) abzielen. Im Fall von *Schleswig-Holstein* heißt ein Qualitätsbereich im Orientierungsrahmen „Inklusion und Umgang mit Heterogenität“ (Ministerium für Schule und Berufsbildung Schleswig-Holstein 2016, S. 12). Unter diesen Qualitätsbereich fallen unter anderem die Diagnose von Lernständen sowie Förderung und Leistungsrückmeldungen. Hierbei liegt es nahe, dass sich Inklusion auf die Gruppe von Schüler_innen mit Behinderung bezieht. So müsste Inklusion als Begriff nicht explizit aufgeführt werden, wenn es die gleiche Bedeutung hätte wie „Umgang mit Heterogenität“ bzw. vice versa. Jedoch lässt sich eine eindeutige Definition von Inklusion daraus nicht ablesen, da in dem Qualitätsbereich Schüler_innen mit diagnostizierter Behinderung nicht explizit genannt werden. Demnach könnte daraus auch ein weiter Inklusionsbegriff abgeleitet werden. Im Fall von *Nordrhein-Westfalen* wird vom „Gemeinsamen Lernen von Schülerinnen und Schülern mit und ohne sonderpädagogischen Unterstützungsbedarf“ gesprochen (NRW_Kommentierung UBB).

Das Fehlen einer expliziten Definition von Inklusion in den Instrumenten und Beschreibungen von guter Schule ist nicht ungewöhnlich, da nur wenige Schulinspektionen über ein Glossar verfügen und ausführliche Definitionen von pädagogischen Konzepten in Fragebögen generell eher eine Ausnahme darstellen. Fehlende Definitionen können jedoch problematisch sein, wenn über ein Konzept – wie in diesem Falle über Inklusion – unterschiedliche Deutungen existieren. Beispielsweise wird in Bayern im Lehrerfragebogen erfasst, ob sich die Lehrkräfte an der Schule regelmäßig über Inklusion informieren. Auch wenn Bayern im Handbuch eine

Definition von Inklusion aufführt, fehlt es im konkreten Lehrerfragebogen an einer expliziten Definition. Dies ist im Hinblick auf die Validität des Instruments kritisch zu sehen, da es sein kann, dass das Antwortverhalten der befragten Personen davon beeinflusst wird, was diese unter Inklusion verstehen (vgl. Gasterstädt und Urban 2016 in Bezug auf Einstellungsforschung zu Inklusion).

Zusätzlich zu den oben aufgeführten Bundesländern gibt es vier Fälle, in denen Inklusion als Begriff in den Instrumenten nicht verwendet wird: Hessen, Saarland, Sachsen-Anhalt sowie Brandenburg.

- Im Fall von *Hessen* taucht Inklusion zwar nicht in den Instrumenten auf, wird aber im Hessischen Referenzrahmen aufgeführt, z. B. „Die Schule fördert die Schulgemeinschaft und die soziale Integration – insbesondere das Zusammenleben der Kulturen sowie die Gleichberechtigung der Geschlechter und der Menschen mit Behinderungen im Sinne einer Weiterentwicklung zur inklusiven Schule“ (Hessisches Kultusministerium 2011, S. 71). Des Weiteren wird auf den Seiten des Kultusministeriums auf die „Checkliste Inklusion“ verwiesen. Diese basiert auf einzelnen Indikatoren des Hessischen Referenzrahmens, die auch der Hessischen Schulinspektion als Orientierung dienen. Jedoch hat das Hessische Projektbüro Inklusion bei der Erstellung der Checkliste Ergänzungen zu den bestehenden Indikatoren vorgenommen. Beispielsweise gibt es im Hessischen Referenzrahmen den Indikator „Im Schulprogramm sind Entwicklungsvorhaben formuliert, an denen zielgerichtet gearbeitet wird“ (ebd., S. 42). Daran anknüpfend enthält die Checkliste Inklusion „Im Schulprogramm ist *die inklusive Schule* als Entwicklungsvorhaben formuliert, an denen zielgerichtet gearbeitet wird“ (Projektbüro Inklusion 2014). Diese Ergänzungen des Projektbüros wurden jedoch von Seiten der Inspektion nicht in die Instrumente übernommen. Demnach taucht Inklusion als Begriff in den analysierten Instrumenten nicht auf.
- Im Fall von *Sachsen-Anhalt* taucht der Begriff in den Messinstrumenten auch nicht auf, jedoch gibt es im Orientierungsrahmen den Indikator „Die Schule ist offen für Inklusion“ (Landesinstitut für Schulqualität und Lehrerbildung Sachsen-Anhalt 2013, S. 5). Es gibt jedoch keine Definition von Inklusion.
- Im Fall des *Saarlands* findet sich keine Nennung des Inklusionsbegriffs im Orientierungsrahmen.
- In *Brandenburg* gab es keinen Orientierungsrahmen, der zwischen 2009 und Oktober 2015 (ausgewählter Zeitraum der Datenanalyse → Kapitel 6) erschienen ist. In der Version von 2016 (vgl. Ministerium für Bildung, Jugend und Sport Brandenburg 2016) wird Inklusion als Begriff genannt, aber ohne nähere Definition.

Zusammenfassend zeigt sich, dass die Bundesländer sich darin unterscheiden, ob sie Inklusion explizit als Begriff verwenden bzw. ob sie diesen näher definieren (für eine Übersicht → Tabelle 7.2). In den Fällen, in denen Inklusion explizit definiert wird, werden in Bezug auf die Zielgruppe unterschiedliche Verständnisse von Inklusion sichtbar. Die Nicht-Verwendung des Inklusionsbegriff kann, unabhängig davon, ob es intendiert ist oder nicht, diverse Implikationen nach sich ziehen: Durch die Entwicklung von Instrumenten in Bezug auf Inklusion aber gleichzeitige Nicht-Verwendung des Inklusionsbegriffs kann das Konzept von Seiten der Schulinspektion als normative Forderung implementiert werden, ohne dass negative Assoziationen mit den Instrumenten geweckt werden. Dies kann eine Strategie sein, wenn Inklusion als Begriff von schulischen Vertretern mit negativen Assoziationen verknüpft wird. Des Weiteren hat die Nicht-Nennung zur Folge, dass Schulen durch die Instrumente der Schulinspektionen keine

Tab. 7.2: Definition von Inklusion in den Instrumenten

	Bundesländer
Explizite Definition von Inklusion ist vorhanden	Bayern, Berlin, Hamburg und Baden-Württemberg
Verwendung des Inklusionsbegriffs, aber keine explizite Definition vorhanden. (Hinweise auf Inklusionsdefinition ergeben sich aus dem Zusammenhang)	Bund-Länder-Inspektion; Niedersachsen und Schleswig-Holstein, Nordrhein-Westfalen
Inklusion wurde als Begriff in den analysierten Instrumenten nicht gefunden	Brandenburg, Hessen, Saarland sowie Sachsen-Anhalt

Informationen darüber erhalten, welche konkreten Implikationen die Umsetzung von Inklusion als politische Vorgabe für sie hat. Demnach werden Schulinspektionen in ihrer normgebenden Funktion (bewusst oder unbewusst) nicht tätig (nähere Ausführung → Kapitel 8).

7.2.2 Instrumente in Bezug auf Inklusion

Im Folgenden geht es nun darum, wie Inklusion im Rahmen der Instrumente operationalisiert wurde. Hierbei geht es nur um diejenigen Bundesländer, in denen Instrumente zu Inklusion identifizierbar sind.

Integration der Instrumente in das bestehende Instrumentenset: Obwohl in vier Bundesländern (BB; HE; SL; ST) keine inklusionsorientierten Instrumente in der Analyse identifiziert wurden, kann es sein, dass Inklusion als normative Forderung in den Instrumenten der jeweiligen Bundesländer berücksichtigt wird. So zeigte sich in Abschnitt 7.1 auch in diesen vier Fällen, dass Aspekte von Inklusion in den Instrumenten thematisiert werden. Dennoch fielen einige der oben aufgeführten Bundesländer bei einzelnen Aspekten auf: So zeigte sich, dass „Behinderung“ bzw. „sonderpädagogischer Förderbedarf“ als Differenzlinie im Fall von Brandenburg, Hessen und des Saarlands (sonst nur in Hamburg) nicht explizit genannt wurden. Im Hinblick auf die Integration der Instrumente zu Inklusion in das bestehende Instrumentenset lässt sich zwischen drei Formen unterscheiden (für einen Überblick → Tab. 7.3):

1. Indikatoren zu Inklusion bzw. Erläuterungen stellen eine Ergänzung zum bestehenden Instrumentenset dar und finden nur in bestimmten Situationen Anwendung;
2. Indikatoren zu Inklusion sind in das bestehende Instrumentenset integriert und finden sich als Querschnitt in verschiedenen Qualitätsbereichen wieder;
3. Indikatoren zu Inklusion stellen im bestehenden Instrumentenset einen eigenständigen Qualitätsbereich dar. Es variiert zwischen den Bundesländern, wann dieser Qualitätsbereich Anwendung findet.

¹ Auch diese Instrumente zum gemeinsamen Unterricht werden als Instrumente zu Inklusion gezählt, weil in der expliziten Definition der Bundesländer unter Inklusion das Konzept „gemeinsamer Unterricht“ verstanden wird.

Tab. 7.3: Integration der Instrumente im Instrumentenset

	Art der Instrumente	Bundesländer
Instrumente zu Inklusion sind als solche gekennzeichnet	a) Zusätzliche Indikatoren bzw. Erläuterungen im Fall von gemeinsamem Unterricht von Schüler_innen mit und ohne Behinderung	Baden-Württemberg (kursive Indikatoren) ¹ Niedersachsen Nordrhein-Westfalen
	b) Integration in das bestehende Instrumentenset	Bayern ¹ Hamburg
	c) Eigenständiger Qualitätsbereich im bestehenden Instrumentenset (Wahl oder Pflicht)	Berlin Schleswig-Holstein Bund-Länder-Inspektion
Keine Instrumente vorhanden, die als Instrumente in Bezug auf Inklusion identifiziert werden können.	—	Brandenburg Hessen Saarland Sachsen-Anhalt

Bei der Beschreibung der Instrumente in Bezug auf Inklusion sind zwei Merkmale von Relevanz: Anwendung der Instrumente sowie Grad der Verpflichtung.

- Als ein Merkmal ist zu nennen, für welche Schulen die Instrumente zu Inklusion gelten. Im Fall von *Bayern*, *Hamburg* und *Schleswig-Holstein* können alle Schulen – unabhängig von ihrer Schülerschaft – zur Umsetzung von Inklusion evaluiert werden. Es gibt jedoch auch Bundesländer, in denen die Instrumente nur für diejenigen Schulen von Relevanz sind, die auch Schüler_innen mit diagnostizierter Behinderung unterrichten (BW, BE, NRW, NI, BLI). Beispielsweise gibt es im Fall von *Baden-Württemberg* zusätzliche Indikatoren, die nur an denjenigen Schulen angewandt werden, wo Schüler_innen mit sonderpädagogischem Bildungsangebot unterrichtet werden.³ Eine ähnliches Vorgehen zeigt sich im Fall von *Nordrhein-Westfalen* mit dem so genannten „I-Schalter“⁴: Nur diejenigen Schulen, wo „Gemeinsames Lernen von Schülerinnen und Schülern mit und ohne sonderpädagogischen Unterstützungsbedarf“ stattfindet, erhalten die jeweiligen Indikatoren aus dem Unterrichtsbeobachtungsbogen (siehe UBB_NRW). Im Fall von *Niedersachsen* gibt es zwar keine zusätzlichen Indikatoren bei der Umsetzung von gemeinsamem Unterricht, jedoch Erläuterungen bei bestehenden Indikatoren, z. B. bei dem Indikator „Die Schülerinnen und Schüler organisieren den Arbeitsprozess eigenverantwortlich“ (UBB_NI) gibt es folgende Ergänzung: „Inklusi-

3 „Sonderpädagogisches Bildungsangebot“ ist im Fall von Baden-Württemberg die bundeslandspezifische Bezeichnung von sonderpädagogischem Förderbedarf, d. h. die Indikatoren zählen nur für diese Gruppe und andere Schüler_innen, die auch unter den Schutz der BRK fallen, werden nicht explizit berücksichtigt (vgl. Piezunka et al. 2016).

4 Der I-Schalter wurde beim nächsten Zyklus wieder abgeschafft, weil er in der Praxis aufgrund unterschiedlicher Anwendungen von Seiten der Inspektor_innen nicht funktioniert hat

on/FÖS GE und KME: Das Maß der erforderlichen Unterstützung ist abhängig vom Grad der Beeinträchtigung“ (UBB_NI).

Im Fall von *Berlin* zielen die Instrumente primär auf die „entsprechenden bezirklichen Schwerpunktschulen“ (Berliner Senatsverwaltung für Bildung, Jugend und Wissenschaft 2012, S. 7) – die so genannten INKA-Schulen⁵ – ab, wo Schüler_innen mit sonderpädagogischem Förderbedarf unterrichtet werden. Die Instrumente waren für diese Schulen verpflichtend und für andere Schulen frei wählbar. Demnach zielen auch im Fall von Berlin die Instrumente primär auf Schulen ab, wo Schüler_innen mit diagnostizierter Behinderung unterrichtet werden. Dies steht insofern im Widerspruch zu der von Berlin verwendeten Definition von Inklusion (siehe oben), wonach unter Inklusion das Ermöglichen von Teilhabe für alle Schüler_innen verstanden wird. Ausgehend von ihrer eigenen Definition dürften Instrumente zu Inklusion nicht nur an bestimmten Schulen Anwendung finden. Im Fall der *Bund-Länder-Inspektion* fehlt es – wie bereits oben dargestellt – an einer klaren Definition von Inklusion in den Instrumenten.

- **Inklusion ist für die Schulen freiwillig oder verpflichtend:** In Bezug auf dieses Merkmal muss prinzipiell zwischen den Schulinspektionen unterschieden werden, die Schulen einen Wahlbereich anbieten, und solchen Inspektionen, wo alle Instrumente verpflichtend sind. Auch wenn es Wahlbereiche gibt, bedeutet dies nicht automatisch, dass die Instrumente zu Inklusion in den jeweiligen Bundesländern frei wählbar sind. Im Fall von *Schleswig-Holstein* sind die Instrumente zu Inklusion Teil des Pflichtbereichs. Im Fall von *Berlin* sind die Instrumente zu Inklusion aus dem „Qualitätsbereich E6“ für die so genannten INKA-Schulen verpflichtend und für andere Schulen frei wählbar. Dagegen werden im Fall der *Bund-Länder-Inspektion* nur diejenigen Schulen zu Inklusion evaluiert, die „auf dem Feld der Inklusion [das, ap] bereits Erreichte [...] gerne präsentieren möchte[n, ap]. Dies soll ein Anreiz für Schulen sein, sich in diesem Bereich weiterzuentwickeln“ (Bundesverwaltungsamt 2014, 15f.).

In den anderen Fällen, die über eindeutig identifizierbare Instrumente zu Inklusion verfügen (BW, HH, NRW), sind alle Instrumente für die Schulen verpflichtend. Hierbei ist jedoch zu berücksichtigen, dass im Fall von Baden-Württemberg und Nordrhein-Westfalen die Instrumente zu Inklusion trotzdem nur bei denjenigen Schulen Anwendung finden, wo Schüler_innen mit diagnostizierter Behinderung unterrichtet werden.

Inhalte der Instrumente zu Inklusion: Die Bundesländer unterscheiden sich nicht nur in der Integration der Instrumente zu Inklusion in das bestehende Instrumentenset, sondern auch in den Inhalten. Generell zeigte sich bei der Extraktion, dass die Inhalte der Instrumente grob folgenden Themenbereiche zugeordnet werden können:

- **Situation von Schüler_innen mit Behinderung:** In vielen Bundesländern gelten diejenigen Instrumente als inklusionsorientiert, die die Situation von Schüler_innen mit diagnostizierter Behinderung bzw. sonderpädagogischem Förderbedarf in den Blick nehmen (ÜI_BLI; ÜI_BW; EFB_BE; UBB_NRW; UBB_NI). Hierbei zeigt sich, dass eine Vielzahl von unterschiedlichen Aspekten in den einzelnen Bundesländern erfasst wird, z. B. ob Schüler_innen mit Behinderung an der Schule aufgenommen werden (ÜI_BLI). Es wird auch überprüft, ob sie die für sie

⁵ Die INKA-Schulen sind Schulen mit besonderer Ressourcenausstattung, wo Kinder mit sonderpädagogischem Förderbedarf in den Bereichen „Lernen“ und „Emotional-Soziale Entwicklung“ unterrichtet werden – ohne im klassischen Sinne diagnostiziert zu werden.

notwendigen Kommunikationsmittel erhalten oder inwiefern eine Lernatmosphäre geschaffen wird, die auch Schüler_innen mit diagnostizierter Behinderung integriert (UBB_NRW; ÜI_BW). Durch die große Varianz an Indikatoren wird deutlich, dass es an einem einheitlichen Set an Indikatoren mangelt.

- **Umgang mit heterogenen Gruppen:** Wie bereits oben erwähnt, verstehen einige Bundesländer unter Inklusion nicht (nur) die Integration von Schüler_innen mit Behinderung, sondern den Umgang mit heterogenen Lerngruppen (BE, HH, SH, BLI). In diesem Fall gelten als explizit gekennzeichnete Indikatoren zu Inklusion, inwiefern z. B. auf sprachliche oder kulturelle Unterschiede eingegangen (EFB_HH) oder ob im Rahmen der Unterrichtsgestaltung differenziert wird (ÜI_BE; ÜI_SH; ÜI_BLI).
- **Auseinandersetzung im Rahmen von Schulentwicklungsprozessen:** Hierbei geht es um Indikatoren, die Schulentwicklungsprozesse bzw. Organisationsprozesse an Schulen thematisieren, z. B. Lehrerkonferenzen, Schulprogramm usw. In Bezug auf diesen Themenbereich lassen sich zwei Arten von Indikatoren unterscheiden:
 - a. Im Fall von *Bayern* thematisieren die Indikatoren, ob eine Auseinandersetzung zu Inklusion stattgefunden hat. Sie sind dadurch gekennzeichnet, dass die theoretische Auseinandersetzung und weniger die konkrete Implementation in der Praxis eine Rolle spielt, z. B. „Die Schule setzt sich mit der Frage auseinander, wie die Teilhabe von Schülern mit besonderen physischen und/oder psychischen Voraussetzungen ermöglicht werden kann“ (ÜI_BY; siehe auch LFB_BY).
 - b. Dagegen gibt es in anderen Bundesländern Indikatoren, die in Bezug auf Schulentwicklungsprozesse überprüfen, inwiefern die Umsetzung von Inklusion in der Praxis bereits implementiert ist, z. B. „Die inklusive Pädagogik ist ein Schwerpunkt im Schulprogramm“ (ÜI_BE; siehe auch ÜI_SH; LFB_HH; ÜI_BLI). Demnach geht es in diesen Fällen weniger um die Vorbereitung in Hinblick auf Inklusion, sondern um die bereits stattgefundene Umsetzung.
- **Vorhandensein von Ressourcen/ Unterstützung durch Kooperationen:** Hierbei geht es um die Frage, ob die Schule bzw. einzelne Lehrkräfte durch andere Akteure Unterstützung erhalten bzw. auf bestehende Ressourcen zurückgreifen können, z. B. „Die Schulleitung unterstützt im Zusammenhang mit der inklusiven Schulentwicklung eine weitreichende Zusammenarbeit mit verschiedenen Ressourcen“ (LFB_HH; siehe auch ÜI_NI; LFB_BY).
- **Qualifikation des Personals:** Im Fall von *Bayern* (LFB_BY) und der *Bund-Länder-Inspektion* (ÜI_BLI) gibt es Indikatoren zur Qualifikation von Lehrkräften, z. B. „Ich habe schon Fortbildungen besucht, die mir dabei helfen, Schüler mit Behinderungen bzw. sonderpädagogischem Förderbedarf zu unterrichten“ (LFB_BY).
- **Einstellung zu Inklusion bzw. zum gemeinsamen Unterricht von Schüler_innen mit und ohne sonderpädagogischen Förderbedarf:** Wie bereits in Abschnitt 7.1.4 thematisiert, stellt das Bundesland *Bayern* eine Ausnahme dar, da Einstellungen zum gemeinsamen Unterricht erfasst werden. Beispielsweise gibt es im Lehrerfragebogen Indikatoren zur Selbstwirksamkeit: „Es gelingt mir, Schüler mit Behinderungen bzw. mit sonderpädagogischem Förderbedarf in die Klassengemeinschaft einzubinden“ (LFB_BY) oder „Ich weiß, wie ich meinen Unterricht anpassen muss, um Schüler mit Behinderungen bzw. sonderpädagogischem Förderbedarf unterrichten zu können“ (LFB_BY).

Zusammenfassend lässt sich beobachten, dass die Bundesländer sich darin unterscheiden, welche Themenbereiche angesprochen werden. Auch innerhalb der Themenbereiche werden

unterschiedliche Indikatoren in Bezug auf den Grad der Anforderungen gewählt. Folgende Tendenzen lassen sich herausarbeiten:

- Einige Bundesländer verstehen (fast überwiegend) unter Instrumenten zu Inklusion solche, die die Situation von Schüler_innen mit diagnostizierter Behinderung in den Blick nehmen: *Niedersachsen, Nordrhein-Westfalen und Baden-Württemberg*. Dabei sind die Instrumente zu Inklusion Zusatzindikatoren im UBB, die Anwendung finden, wenn sich im Klassenraum Schüler_innen mit diagnostizierter Behinderung befinden.
- Folgende Bundesländer nehmen noch weitere Aspekte in den Blick, wie Umgang mit heterogenen Lerngruppen sowie Schulentwicklung: *Hamburg, Schleswig-Holstein, Berlin und Bund-Länder-Inspektion*. Im Fall von *Berlin* ist jedoch zu berücksichtigen, dass die Instrumente zu Inklusion sich inhaltlich nicht nur auf die Situation von Schüler_innen mit diagnostizierter Behinderung beziehen, aber dass die Instrumente für Inklusion nur für diejenigen Schulen verpflichtend sind, die solche Schüler_innen unterrichten.
- *Bayern* weicht von den anderen Bundesländern ab, da die Indikatoren primär darauf abzielen, inwiefern sich die Schulen bzw. die Lehrkräfte vorbereitet fühlen und wie sich einzelne Akteure zum gemeinsamen Unterricht positionieren.

7.2.3 Zwischenfazit: Verständnisse von Inklusion in den Instrumenten

Nachdem in Abschnitt 7.1 analysiert wurde, inwiefern Inklusion als normative Forderung in den Instrumenten berücksichtigt wurde, wurde im nächsten Schritt betrachtet, wie Inklusion als Begriff in den Instrumenten der Schulinspektionen operationalisiert wurde. Hierfür wurde näher betrachtet, ob und wie Inklusion in das bestehende Set an Instrumenten integriert wurde und welche Themenbereiche durch die Instrumente angesprochen wurden. Unter Berücksichtigung von drei Aspekten ließen sich grob vier Typen herausarbeiten (für eine Übersicht → Tab. 7.4).

Zusammenfassend zeigt sich, dass in vier Fällen unter Inklusion der gemeinsame Unterricht von Schüler_innen mit und ohne diagnostizierte Behinderung verstanden wird (NRW, NI, BW, BY). *Hamburg* und *Berlin* stellen eine Ausnahme dar, da sie sich von diesem Inklusionsverständnis distanzieren und explizit auf die Salamanca-Konferenz (EFB_HH) bzw. im Fall von *Berlin* auf die in der BRK genannte Ziele wie Diskriminierungsfreiheit (ÜI_BE) verweisen. Im Fall von *Schleswig-Holstein* und der *Bund-Länder-Inspektion* geben die Inhalte der Instrumente Hinweise darauf, dass es nicht nur um die Integration von Schüler_innen mit diagnostizierter Behinderung geht. Darüber hinaus gibt es vier Fälle (BB, SL, ST, HE), in denen Inklusion als Konzept überhaupt nicht vorkommt. Hierbei zeigen sich zum Teil Übereinstimmungen zum ersten Teil der Analyse (→ Abschnitt 7.1): So ergab sich hier, dass im Fall von Brandenburg und des Saarlands viele Aspekte in Bezug auf das hier verwendete Verständnis von Inklusion kaum thematisiert wurden. Diese Beobachtung trifft jedoch nicht auf Hessen und Sachsen-Anhalt zu.

In Bezug auf Typ C ist es überraschend, dass die Indikatoren zu Inklusion nur in Schulen Anwendung finden (im Fall der Bund-Länder-Inspektion unklar formuliert), wo Schüler_innen mit Behinderung unterrichtet werden, obwohl die Indikatoren nicht nur auf diese Gruppe, sondern auch auf den Umgang mit Heterogenität im Allgemeinen abzielen. So könnte man argumentieren, dass diese Aspekte auch alle anderen Schulen betreffen. Bei Typ D (Bayern) wird ein enges Inklusionsverständnis sichtbar, aber die Instrumente zu Inklusion gelten trotzdem für alle Schulen. Dies ist insofern nicht überraschend, als die Instrumente zu Inklusion nicht die Situation von einzelnen Schüler_innen mit Behinderung in den Fokus nehmen, sondern

Tab. 7.4: Messungen von Inklusion

	Bundesländer	Inklusionsverständnis und Inhalte der Instrumente	Anwendung der Instrumente	Integration in das bestehende Instrumentenset
Typ A	HH, SH	Weiter Inklusionsbegriff Indikatoren thematisieren Umgang mit Heterogenität	Alle Schulen	Indikatoren zu Inklusion gehören zum bestehenden Instrumentenset
Typ B	NRW, NI, BW	Enger Inklusionsbegriff Bei den Indikatoren bzw. Erläuterungen steht die Situation von Schüler_innen mit diag. Behinderung im Vordergrund	Regelschulen, die Schüler_innen mit diagnostizierter Behinderung unterrichten	Zusätzliche Indikatoren bzw. Erläuterungen
Typ C	BLI, BE	Weiter Inklusionsbegriff Bei den Indikatoren geht es unter anderem um den Umgang mit heterogenen Gruppen. Im Fall der BLI auch um die Situation von Schüler_innen mit diag. Behinderung.	Regelschulen, die Schüler_innen mit diagnostizierter Behinderung unterrichten	Eigener Qualitätsbereich
Typ D	BY	Enger Inklusionsbegriff Indikatoren erfassen Auseinandersetzung mit Inklusion sowie Haltungen.	Alle Schulen	Indikatoren zu Inklusion gehören zum bestehenden Instrumentenset

erfassen, inwiefern sich die Schulen bereits mit Inklusion auseinandergesetzt haben und welche Haltungen in Bezug auf Inklusion bestehen.

Des Weiteren ist es auffällig, dass vorhandene Daten zur Umsetzung von Art. 24 BRK (nähere Ausführung → Abschnitt 2.3) – wie die Exklusionsquote – keine Hinweise darauf geben, dass es einen Zusammenhang zwischen der Umsetzung im jeweiligen Bundesland und der Berücksichtigung von Inklusion in den Instrumenten gibt: So weist Sachsen-Anhalt im bundesdeutschen Vergleich eine recht hohe Exklusionsquote auf (vgl. Bertelsmann-Stiftung 2014, S. 11). Dagegen sind die Werte von Brandenburg, Hessen und dem Saarland sehr nah am bundesdeutschen Durchschnitt (Exklusionsquote von 4,8 Prozent, siehe ebd., S. 11). Hierbei ist jedoch zu berücksichtigen, dass die Exklusionsquote nur einen Teilaspekt der Umsetzung von Art. 24 BRK und somit auch nur einen kleinen Aspekt des hier verwendeten Inklusionsverständnisses repräsentiert.

Die unterschiedlichen Deutungen haben zur Folge, dass in den jeweiligen Bundesländern die Schulen unterschiedliche Signale erhalten, was die Umsetzung von Inklusion anbelangt. So unterscheiden sich die Deutungen darin, welche inhaltlichen Anforderungen an die Schulen formuliert werden, z. B. Fokus auf Schüler_innen mit zugeschriebener Behinderung oder Umgang mit heterogenen Lerngruppen. Des Weiteren gibt es auch Unterschiede im Grad der Anforderungen, z. B. wenn im Fall von Bayern abgefragt wird, ob sich die Schulen bereits mit dem Thema auseinandergesetzt haben, während in anderen Bundesländern bereits die konkrete Implementation abgefragt wird. Der Grad der Anforderungen spiegelt sich auch in der Integration der Indikatoren zu Inklusion in das bestehende Instrumentenenset wider, z. B. ob und für wen die jeweiligen Indikatoren verpflichtend sind.

Der Fokus auf Schüler_innen mit zugeschriebener Behinderung in vielen Bundesländern ist insofern nicht überraschend, als Inklusion als Konzept durch die BRK Eingang in die öffentliche Debatte gefunden hat. Hierbei ist jedoch zu ergänzen, dass in einigen Bundesländern (z. B. Baden-Württemberg) nur eine Teilgruppe der Personen, die unter den Schutz der BRK fallen (vgl. Piezunka et al. 2016), in den Blick genommen wird. Des Weiteren ist auch keine systematische Ableitung der Anforderungen erkennbar, die sich aus Art. 24 BRK ergibt (vgl. Piezunka et al. 2018).

7.3 Inklusionsdeutungen der Instrumenteentwickler_innen

Die Instrumente der Inspektionen stellen das Resultat des Sensemakings auf der intersubjektiven Ebene dar. Im Folgenden geht es nun um die Inklusionsverständnisse von einzelnen Personen, die in den jeweiligen Aushandlungsprozessen involviert waren. Die Verständnisse der Instrumenteentwickler_innen sind das Resultat des Sensemakings auf individueller Ebene. Diese individuellen Verständnisse beeinflussen wiederum die Entwicklung der Instrumente (nähere Ausführung → Kapitel 8).

Aus den Interviews lassen sich drei Deutungen von Inklusion herausarbeiten, die sich insbesondere in der Frage unterscheiden, warum Inklusion in den Instrumenten der Schulinspektion berücksichtigt werden sollte. Des Weiteren spielen bei der Unterscheidung folgende Dimensionen eine Rolle: Definition von Inklusion, Zielgruppe, Beschreibung von inklusiver Schule, Machbarkeit, Grad der Anforderungen sowie Abgrenzung zu dem jeweiligen Verständnis von guter Schule (→ Tabelle 7.5).

7.3.1 Typ I: Inklusion als Anspruch auf Förderung

Inklusion wird als ein Anspruch auf Unterstützung angesehen, der für Schüler_innen mit Behinderung (= Zielgruppe) an der Regelschule besteht: *„Wir haben ja hier ne klare Definition, wann Schüler inklusiv beschult werden an einer allgemeinen Schule durch eine Sonderschullehrkraft. Und genau auf diesen speziellen Fall, wo auch dieser Anspruch gegeben ist“* (Int. 2_24). Dahinter steht die Idee, dass Schüler_innen, die nicht mehr an der Förderschule, sondern nun an der

Tab. 7.5: Deutungen der Instrumenteentwickler_innen

	Typ I: Inklusion als Anspruch auf Förderung	Typ II: Inklusion als Herausforderung	Typ III: Inklusion als Teilaspekt von „guter“ Schule
Begründung für aktives Handeln durch die Schulinspektion	Es ist sicherzustellen, dass bestimmte Schüler_innen Anspruch auf Förderung haben.	Schulen müssen sich auf dem Weg machen, Inklusion umzusetzen und die Inspektionen beobachten, inwiefern sie aktiv werden.	Die Inspektionen messen Schulqualität und der gelungene Umgang mit heterogenen Lerngruppen ist für sie eine Beschreibung von guter Schule.
Definition von Inklusion	Gemeinsame Beschulung von Schüler_innen mit und ohne Behinderung	Gemeinsame Beschulung von Schüler_innen mit und ohne Behinderung	Gelungener Umgang mit heterogenen Lerngruppen
Zielgruppe	Fokus auf Schüler_innen mit Behinderung	Fokus auf Schüler_innen mit Behinderung	Heterogene Lerngruppe
Beschreibung von inklusiver Schule	Regelschulen mit Schüler_innen mit Behinderung	Regelschulen mit Schüler_innen mit Behinderung	Alle Schulen sollten sich bemühen, inklusiv zu sein.
Machbarkeit	Umsetzung von Inklusion wird als machbar angesehen. Ansprüche, die sich aus der Förderschule ergeben, sollen in der Regelschule aufrechterhalten werden.	Umsetzung von Inklusion stellt eine schwer umsetzbare Herausforderung dar	Alle Schulen sollen sich um den erfolgreichen Umgang mit heterogenen Lerngruppen bemühen.
Grad der Anforderungen, die an Schulen formuliert werden	Mindeststandards ergeben sich aus dem Vergleich mit Förderschulen	Inklusion wird als Prozess gesehen und Überforderung von Schulen soll vermieden werden	<i>Keine Angaben</i>

	Typ I	Typ II	Typ III
Bezug zum Verständnis von „guter“ Schule	Kann einen Teilaspekt von „guter“ Schule darstellen, indem Inklusion als Recht gesehen wird und Schulen „gut“ sind, wenn sie rechtliche Vorgaben einhalten.	Herangehensweise kann implizieren, dass die Umsetzung von Inklusion skeptisch gesehen wird. Demnach wäre Inklusion nicht notwendigerweise eine Beschreibung von guter Schule.	Inklusion ist ein Teilaspekt von guter Schule

Regelschule beschult werden, weiterhin die gleichen Formen von Unterstützung erhalten sollen, die sie an der Förderschule erhalten haben: *„Wenn Schule sich auf den Weg macht, ist toll, aber dann muss sie auch den Bedarf an Förderung [anonymisiert, ap] einlösen. Weil das Kind ja, wenn es ja an einer Förderschule wäre, dann hätte es ja auch die Garantie, dass der Bedarf an Förderung [anonymisiert, ap] eingelöst wird“* (Int. 1_26). Demnach ist die Befürchtung vorhanden, dass die Schüler_innen an der Regelschule nicht mehr die Unterstützung erhalten, die sie benötigen (Int. 3_2). Aus dieser Perspektive heraus stellt sich nicht die Frage der Machbarkeit, sondern es ist außer Frage, dass die Förderung, die sich aus dem schuladministrativen Etikett „Förderbedarf“ ergibt, auch in der Regelschule eingelöst werden muss: *„Dann hat jede Schule, die Schüler mit Behinderung [anonymisiert, ap] hat letztendlich diese Kriterien [anonymisiert, ap] zu erfüllen – gleich, ob das fünf Schüler sind oder ein Schüler ist – weil genau dieser Schüler, der diesen Anspruch hat, muss diesen Anspruch an dieser Schule – die Schule muss das einlösen“* (Int. 9). Demnach werden die Ansprüche als Mindeststandards angesehen, die Schulen erfüllen müssen und demnach auch als machbar angesehen werden.

Als inklusive Schulen gelten dabei nur diejenigen Schulen, die Schüler_innen mit diesem Status haben: *„Wir wissen ja, dass noch nicht alle Schulen inklusiv unterwegs sind. Es ist der Wunsch und die Erwartung seitens der Landesregierung, dass möglichst viele Schulen solche Schüler aufnehmen und inklusiv beschulen“* (Int. 2_30). Hinter dieser Argumentation steht ein Verständnis von Inklusion als gemeinsame Beschulung von Schüler_innen mit und ohne Behinderung. Daher geht es nicht um die Frage, ob eine inklusive Schule eine gute Schule ist. So macht diese Herangehensweise keine Aussagen über die Zielsetzung von Schule im Allgemeinen. Hierbei wird stattdessen ein Verständnis von guter Schule sichtbar, wonach es darum geht, rechtliche Vorgaben und die Umsetzung von bildungspolitischen Reformen einzuhalten. Eine direkte Verknüpfung zwischen „guter“ Schule und Inklusion wird nicht sichtbar.

7.3.2 Typ II: Inklusion als Herausforderung für Schulen

Zwischen Typ I und II gibt es insofern Überschneidungen, als beide unter Inklusion den gemeinsamen Unterricht von Schüler_innen mit und ohne zugeschriebene Behinderung verstehen. Sie unterscheiden sich jedoch in der Beantwortung der Frage, warum die Schulinspektion tätig werden sollte. So wird bei Typ II von Seiten der Schulinspektion Inklusion als eine Herausforderung bzw. Belastung verstanden, die Schulen nun meistern müssten: *„Ich meine, dass eben die Bundesregierung sich verpflichtet hat, vielleicht diese Konvention eben umzusetzen. Und ich glaube, also dadurch ist das sozusagen die Rechtsnorm geworden in Deutschland“* (Int. 8).

Ausgehend von der Definition von Inklusion werden bei Typ II – wie bei Typ I – unter inklusiven Schulen die Regelschulen verstanden, die auch Schüler_innen mit Behinderung unterrichten. Die gemeinsame Beschulung als politische Maßnahme wird von den Vertreter_innen dieses Typs als eine belastende Herausforderung wahrgenommen, die aktuell für die Schulen besteht. Dies schließt jedoch nicht aus, dass sich das Verständnis von Inklusion auf lange Sicht verändern kann: *„An allen Ecken und Enden wird daran gearbeitet, aber man kann natürlich jetzt net einen ganz weiten Begriff von heute auf morgen umsetzen. Das wird sich über die Jahre hin verändern“* (Int. 17_40).

Die Unterscheidung zwischen Typ I und Typ II zeigt sich bei den Teilaspekten Machbarkeit und Art der Messung: So sieht Typ II Inklusion als eine Maßnahme an, die Schulen nun umsetzen müssen, aber die jeweiligen Interviewpartner sehen die Umsetzung der Maßnahme als langwierigen und anspruchsvollen Prozess an: *„Und das Inklusion mit ganz kleinen Schritten anfängt und*

wir das dann von Schulen auch gar nicht die Umsetzung eines großen Gesamtkonzepts jetzt erwarten“ (Int. 8). Dies spiegelt sich in Bezug auf die Anforderungen insofern wider, als – nach der Perspektive der Instrumentenentwickler_innen – niedrige Anforderungen formuliert werden: „Also wir haben ja ein sehr sehr niederschwelliges Maßstab angelegt – der wird sich vielleicht über die Jahre hinweg mal verändern und härter werden“ (Int. 17_47). Die Herangehensweise „Inklusion als Herausforderung“ geht bei manchen Interviewpartnern auch mit einer kritischen Position gegenüber der Frage einher, ob die gemeinsame Beschulung überhaupt gelingen kann: „meine persönliche Meinung ist, dass was Hattie erwartet, was auch jetzt erwartet wird im ganzen Kontext von Differenzierung oder gar Individualisierung – Inklusion ist für die meisten Lehrer im Schulalltag nur schwer erreichbar, nur unbefriedigend erreichbar“ (Int. 8). Demnach wird die Frage nach Schulqualität auch folgendermaßen beantwortet: „Also ne gute Schule hängt nicht an der Frage, ob Inklusion stattfindet oder nicht“ (Int. 18_29). Die Berücksichtigung von Inklusion in den Instrumenten der Schulinspektion lässt sich demnach nicht damit begründen, dass Inklusion einen Teilaspekt von guter Schule darstellt. Darüber hinaus geht es – im Gegensatz zu Typ I – nicht darum, sicherzustellen, dass bestimmte Mindeststandards gegeben sind: „Aber es ist jetzt kein Anspruch und wie weit sind sie in Inklusion. Das ist ja nicht unsere Formulierung, sondern Schulentwicklungsprozesse und der Bezug zur schulischen Situation, zu den Notwendigkeiten aus Schülerschaft, Kollegium ähm auch Potenzialen zum Beispiel im Lehrerkollegium. Wenn man sagt man hat bestimmte Standarderwartungen, dann ist das nicht unsere Sichtweise“ (Int. 5_37). Stattdessen steht die Frage im Mittelpunkt, wie die Schulen mit der Umsetzung von Inklusion umgehen bzw. ob sie sich bereits damit beschäftigen. Das heißt Inklusion wird als ein Reformprozess verstanden, der von den Schulen nun schrittweise umgesetzt werden soll bzw. muss. Dabei wird die Schulinspektion in ihrer Funktion als Beobachterin und Impulsgeberin gesehen.

7.3.3 Typ III: Inklusion als Teilaspekt von guter Schule

Bei Typ III wird Inklusion als Gelingensbedingung für Schulqualität gesehen. Dabei wird unter Inklusion ein erfolgreicher Umgang mit heterogenen Lerngruppen verstanden: „Eine inklusive Schule wäre eine, die nicht in Ressourcen denkt, sondern in Potenzialen der Schülerinnen und Schüler, die also die einzelnen Lernenden in den Blick nimmt und sich fragt, was braucht diese Gruppe, was braucht diese Person. Und das können unterschiedliche Dinge sein, da geht's auch nicht ohne Gleichbehandlung (...), sondern da geht darum den einzelnen Gruppen und den einzelnen Personen gerecht zu werden“ (Int. 5_31). Es findet eine explizite Verknüpfung zwischen dem Verständnis von Schulqualität und Inklusion statt: „Inklusive Schule wäre für mich jede gute Schule. Wenn - wenn es gelingt mit Vielfalt umzugehen, dann äh- (.) dann gelingt Inklusion.“ (Int. 15). Folgt man dieser Argumentation, dann ist die Aufnahme von Schüler_innen mit „Behinderung“ keine große Herausforderung, wenn der Umgang mit heterogenen Lerngruppen im Allgemeinen bereits gelingt: „Also der inklusive Rahmen ist ja nur die Weiterdenke von kompetenzorientierten und individualisiertem Unterricht. Ja und Schulen, die sich dafür öffnen (.) haben damit überhaupt keine Probleme inklusiv auch tätig zu werden. Inklusion wird hier in Bundesland X [anonymisiert, ap], äh beschränkt gedacht, auf die Umsetzung der Behindertenrechtskonvention“ (Int. 14_30). Dabei grenzt sich dieser Typ explizit von den beiden vorher beschriebenen Typen ab, da nicht nur Schüler_innen mit diagnostizierter Behinderung im Fokus stehen, sondern weitere Differenzlinien von Relevanz sind. Hierbei gibt es zwei Subtypen:

- a) Für manche Instrumenteentwickler_innen werden als relevante Differenzlinien diejenigen angesehen, die im Zusammenhang mit Lernprozessen stehen, z. B. kognitive Fähigkeiten. Demnach liegt hier der Fokus auf der Qualifizierungsfunktion von Schule: *„Inklusion bedeutet ein Bildungssystem zu schaffen, indem jeder Schüler, jede Schülerin lernen kann mit ihren individuellen Voraussetzungen“* (Int. 4).
- b) Für andere Instrumenteentwickler_innen werden nicht nur Differenzlinien, die im Hinblick auf den Lernprozess von Relevanz sind, aufgeführt, sondern noch weitere Differenzlinien, die in Bezug auf Akzeptanz von Vielfalt bzw. soziale Integration eine Rolle spielen können: *„Inklusion kann ja schon bedeuten, dass ich andere Religionen akzeptiere, dass ich andere Hautfarben akzeptiere, dass ich akzeptiere die Vielfältigkeit insgesamt und dass ich darauf versuche einzugehen. Das mitzudenken, das mit zu betrachten“* (Int. 11_24). Demnach geht es für diesen Subtyp nicht nur um die Qualifizierungsfunktion von Schule, sondern weitere Aspekte, wie die soziale Integration von Schüler_innen, werden von den Instrumenteentwickler_innen in den Blick genommen.

Allgemein ist zu berücksichtigen, dass es bei den Textpassagen primär um die jeweiligen Verständnisse von Inklusion und nicht um ihr Verständnis von Schulqualität. Beispielsweise wird der erfolgreiche Umgang mit heterogenen Lerngruppen als ein Indikator für Schulqualität nicht nur bei Typ III genannt: *„bei der Leistungsrückmeldung und auch bei der passenden Förderung – da muss es ja für alle passen und dann sind die dort eigentlich mitbedient und laufen unter Heterogenität, weil da für uns auch kulturelle Unterschiede – ja alles, was es da sonst so gibt, im Bereich Heterogenität dann quasi die gleiche Ebene hat oder eben halt die mit Behinderung [anonymisiert] und das sind dann für uns [dann die, ap], die wir quasi als Inklusion greifen. Weil es da ja auch andere Mittelzuweisungen gibt“* (Int. 1_34). Demnach können auch Vertreter_innen von Typ I und II den gelungenen Umgang mit heterogenen Lerngruppen im Hinblick auf Schulqualität für wichtig erachten, aber sie unterscheiden sich insofern von Typ III, als sie dies nicht mit dem Begriff „Inklusion“ verknüpfen.

Aus dieser Beschreibung von Inklusion bzw. aus der Zielgruppe von Typ III ergibt sich – im Vergleich zu Typ I und II – eine andere Beschreibung von inklusiver Schule: *„Wichtig war uns zunächst mal die Entscheidung, dass wir uns nicht auf diesen schmalen Inklusionsbegriff äh beschränken (...) Und damit der Anspruch auch vorhanden ist, bei allen Schulformen zu gucken, wie sie mit Inklusion, Heterogenität umgehen“* (Int. 16_27). So wird als inklusive Schule eine Schule gesehen, der es gelingt, erfolgreich mit heterogenen Lerngruppen umzugehen. In Bezug auf die Machbarkeit wird Inklusion als ein Prozess und nicht als ein zu erreichender Status definiert. Demnach ist Inklusion machbar, weil Beispiele von Schulen genannt werden, denen der Umgang mit heterogenen Lerngruppen gelingt. Jedoch gibt es in diesem Zusammenhang auch immer wieder Verweise auf die Rahmenbedingungen von Schule, die einen gelungenen Umgang mit Heterogenität erschweren, z. B. das Abkapseln des Gymnasiums (z. B. Int. 14) oder die etikettenabhängige Ressourcenzuweisung (z. B. Int. 5_31).

Wie bereits oben erwähnt, wird Inklusion – im Gegensatz zu den anderen beiden Typen – als ein Teilaspekt von guter Schule gesehen: *„Denn das ist im Grunde – ist uns da was deutlich geworden, was auch der Schulqualitätsrahmen [anonymisiert, ap] von Anfang an klar macht, nämlich: Eine gute Schule ist eine inklusive Schule“* (Int. 13_8).

7.4 Deutungen von Inklusion im Vergleich

Bei der Analyse der Interviews haben sich drei unterschiedliche Begründungen dafür herauskristallisiert, warum Inklusion in den Instrumenten der Schulinspektion berücksichtigt werden sollte und wie dies geschehen soll. Die eben betrachteten Inklusionsdeutungen der Instrumentenentwickler_innen stellen das Ergebnis von Sensemaking-Prozessen auf individueller Ebene dar. Die Instrumente sind dagegen die Folge von Aushandlungsprozessen auf kollektiver Ebene.

Im Folgenden geht es nun um die Frage, inwiefern sich die Verständnisse aus den Interviews (Typ I – III; → Abschnitt 7.3) in den Instrumenten (Typ A – D; → Abschnitt 7.2) wiederfinden:

- Der Typ I „Inklusion als Anspruch auf Förderung“ aus den Interviews würde sich in den Instrumenten sichtbar werden, wenn die jeweiligen Instrumente für Schulen mit Schüler_innen mit diagnostizierter Behinderung verpflichtend sind. Dies ist bei Typ B der Instrumente der Fall.
- Typ II „Inklusion als Herausforderung“ aus den Interviews würde implizieren, dass in den Instrumenten nicht die konkrete Umsetzung überprüft wird, sondern untersucht wird, inwiefern eine Auseinandersetzung mit Inklusion stattgefunden hat bzw. welche Haltungen vorhanden sind. Hier lassen sich beispielsweise Anknüpfungspunkte zu Typ D in den Instrumenten finden.
- In Bezug auf Typ III „Inklusion als Teilaspekt von guter Schule“ in den Interviews gibt es Anknüpfungspunkte zu den Inklusionsdeutungen in den Instrumenten, wo Inklusion weiter gefasst wird (Typ A und C).

Nach der eben genannten Beschreibung müssten Interviewpartner, die Typ III zugeordnet werden, aus Bundesländern kommen, die unter Typ A und C fallen. Diese Verknüpfungen lassen sich nur zum Teil herstellen. Aus Datenschutzgründen darf nicht näher ausgeführt werden, aus welchen Bundesländern die jeweiligen Interviewpartner kommen, aber eine nähere Ausführung dazu, weshalb die Zuordnungen nur begrenzt zutreffen, folgt.

7.5 Exkurs: Konsensuales Verständnis von Schulqualität

Durch die Analyse der Messinstrumente und der Interviews ist deutlich geworden, dass es keinen Konsens darüber gibt, was unter Inklusion verstanden wird. Im Rahmen des folgenden Kapitels (→ Kapitel 8), wird es darum gehen, herauszuarbeiten, welche Rahmenbedingungen die Entwicklung von Instrumenten zu Inklusion geprägt haben. Zuvor soll jedoch noch im Rahmen eines Exkurses auf die Idee eines allgemein anerkannten Verständnisses von Schulqualität näher eingegangen werden. Dies ist als Hintergrundinformation von Relevanz, um die Rechtfertigungsordnungen der Instrumentenentwickler_innen einordnen und verstehen zu können. So zeigt sich als Muster in den Dokumenten sowie in vielen der Interviews, dass bei der Entwicklung von Instrumenten auf ein allgemein anerkanntes Verständnis von Schulqualität zurückgegriffen wird, wenn Bewertungsinstrumente entwickelt werden.

So argumentiert Sachsen-Anhalt: „Bei dem Versuch Qualitätsstandards (...) zu formulieren, gehen die „Verfasser“ des Qualitätsrahmens von der Vorstellung einer Guten Schule aus“ (Landesinstitut für Schulqualität und Lehrerbildung Sachsen-Anhalt 2013, S. 2). Es fehlt jedoch im weiteren Dokument bzw. in anderen Dokumenten an einer expliziten Definition, was unter einer „guten Schule“ verstanden wird. Bei solchen Formulierungen entsteht der Eindruck, dass alle

wüssten, was damit gemeint sei. Des Weiteren beschreibt Bayern den Überarbeitungsprozess der Instrumente zwischen dem ersten und zweiten Zyklus folgendermaßen: „Dabei wurden (...) die eingesetzten Instrumente überarbeitet und weiter objektiviert“ (Bayerisches Staatsministerium für Unterricht und Kultus 2005, S. 8). Des Weiteren wird in Bayern argumentiert:

„Externe Evaluation muss frei sein von subjektiven Theorien. Es ist unwichtig, was einzelne Evaluatoren für wichtig halten. Die Kriterien für die Qualität von Schule und Unterricht müssen dem Stand der jeweiligen wissenschaftlichen Forschung genügen, also valide und intersubjektiv nachvollziehbar sein“ (ebd., S. 10).

Im Handbuch von Niedersachsen steht: „Grundsätzlich wird davon ausgegangen, dass die in der wissenschaftlichen Unterrichtsforschung allgemein anerkannten Kriterien für „guten Unterricht“ schulform- und fachunabhängig gelten“ (Niedersächsisches Landesinstitut für schulische Qualitätsentwicklung 2014, siehe auch für Bayern Bayerisches Staatsministerium für Unterricht und Kultus 2005, S. 4).

Die Argumentation, dass es Konsens sei, was unter guter Schule bzw. unter gutem Unterricht verstanden wird, findet sich auch in vielen Interviews wieder. In diesem Zusammenhang werden verschiedene Bezüge hergestellt, die den Eindruck erwecken, als wenn es ein Verständnis von guter Schule gebe, welches allgemein anerkannt wäre: **Erstens wird argumentiert, dass alle Bundesländer das gleiche machen würden:** „*Ich meine die Kriterien guter Schule – da brauchen sie im Grunde nur einen (...) der 14 Qualitätsrahmen der Länder anzusehen. Die ähneln sich alle sehr – meines Erachtens*“ (Int.18_19; siehe auch 4_20; 5_16; 17_21; 7_10; 9_35; 6; 12_21; 8_5; 8_22). Dagegen argumentieren Döbert et al:

„In Bezug auf die in den Unterrichtsbeobachtungsbögen aufgenommenen Qualitätsaspekte finden zwischen Bundesländern zwar Gemeinsamkeiten, eine tiefergehende Analyse zeigt aber, dass die Konzepte auf der Ebene der einbezogenen Items eine hohe Variation aufweisen. (...) Darüber hinaus finden sich auch bei Items, die sich auf den ersten Blick auf die Erfassung desselben Sachverhaltes beziehen, vielfach unterschiedliche Nuancierungen“ (Döbert et al. 2008, S. 146).

Zweitens, wird darauf verwiesen, dass auf internationaler Ebene Konsens über die Kriterien von Schule bzw. Unterricht herrscht: „*Da kann man sich natürlich wieder an den Kriterien des guten Unterrichts, die international einfach anerkannt sind, orientieren*“ (Int. 11, siehe auch 10_27; 8_9). **Drittens werden Bezüge zur Wissenschaft hergestellt.** So wird argumentiert, dass die eigenen Kriterien aus der Wissenschaft kämen und es dort ein einheitliches Verständnis darüber gebe, welche Kriterien relevant sind: „*Es gibt so vereinzelt Unterschiede, aber letztlich ist der Aufbau schon eigentlich ähnlich. Was ja auch daran liegt, dass alle sagen wir stützen uns auf Forschung, das heißt Forschung wird mit einbezogen*“ (Int. 4_20). Dabei ist bei manchen auch das Bild vorhanden, dass es innerhalb der Forschung Einigkeit darüber gebe, was man unter Schulqualität versteht: „*Na gut, Wissenschaft ist sich einig – da fordern wir das auch ein.*“ (Int. 1_44, siehe auch 1_14; 2_12; 18_18; 7_16) oder „*das ist einfach der generelle, auch wissenschaftlich fundierte Blick auf gute Schule, ob man das jetzt mit Hilbert Meyer oder mit der Hattie-Studie definiert - das ist ja eigentlich relativ egal – die decken sich in den ganz großen Bereichen – da hat man ja nicht zwingend was Neues*“ (Int. 11_8). Die Notwendigkeit, das eigene Handeln über Forschung legitimieren zu können, wird anhand dieses Beispiels deutlich: Als eine Interviewperson bemerkt, dass es bei den Instrumenten an wissenschaftlichen Verweisen fehlt, meint sie: „*Es gibt keine Traktur in weiß. Keine Quelle. Nirgends. (AP: Ja.) Und das finde ich zum Beispiel verwunderlich. Als ob sich das jemand komplett alles ausgedacht hat*“ (Int. 5_7).

In den Interviews gab es zugleich einzelne Positionierungen zu Schulqualität, wonach nicht davon ausgegangen wird, dass es allgemein anerkannt sei, was unter einer guten Schule zu verstehen ist. So beschreibt die Senatorin für Bildung, Jugend und Wissenschaft des Landes Berlin im Vorwort des Handlungsrahmens Schulqualität:

„Die Ansichten darüber, was von einer guten Schule erwartet wird, sind vielfältig und entwickeln sich fortwährend weiter. In der öffentlichen Diskussion zeigt sich, dass die an Schule Beteiligten und Interessierten in der Regel recht klare Vorstellungen haben und je nach Interessenslage unterschiedliche Schwerpunkte setzen“ (Senatsverwaltung für Bildung 2012; siehe auch Ministerium für Bildung, Jugend und Sport des Landes Brandenburg 2008).

Des Weiteren steht im Handbuch der Schulinspektion von Schleswig-Holstein:

„Der <Orientierungsrahmen Schulqualität> beschreibt differenziert, was in Schleswig-Holstein unter einer guten Schule verstanden wird. Die Qualitätsbereiche und Qualitätsmerkmale mit ihren Erläuterungen sind jedoch nicht statisch, da sich die Anforderungen an Schule verändern. Der Orientierungsrahmen erhebt daher nicht den Anspruch, alle Kriterien einer guten Schule umfassend zu beschreiben“ (Ministerium für Schule und Berufsbildung Schleswig-Holstein 2016).

Die Positionierung, dass es keine allgemein anerkannten Kriterien gebe, findet sich vereinzelt auch in den Interviews wider (Int. 9_34; 16_19; 8_42): „*Stellen sie sich ne Blume vor, lassen sie n Chemiker ran, lassen sie n Maler ran, lassen sie nen Pfarrer ran, lassen sie ein verliebten jungen Menschen ran. Jeder schaut das gleiche an und wird zu völlig unterschiedlichen Ergebnissen kommen*“ (Int. 8_42). Des Weiteren wird von anderen Instrumenteentwickler_innen die Forschung nicht als einheitlicher Akteur wahrgenommen, der ein einheitliches Verständnis von guter Schule vermittele: „*Es gibt aus der Forschung verschiedene Ansprüche an Unterricht und auch, dass nach Helmke oder ähm Hilbert Meyer oder was auch immer. Das sind verschiedene Ansprüche, auch in Bezug auf neuere Veröffentlichungen, in Bezug auf kognitive Aktivierung zum Beispiel*“ (Int. 5_27).

Auch wenn sich einzelne Interviewpersonen davon distanzieren, verweisen viele Instrumenteentwickler_innen darauf, dass es ein allgemein anerkanntes Verständnis von guter Schule gebe und dieses als Grundlage genutzt wird, um Instrumente zur Messung von Schulqualität zu entwickeln. Im Folgenden wird deutlich, dass jene Argumentation sich in Bezug auf die Entwicklung von Instrumenten zu Inklusion als ein Teilaspekt von Schulqualität nicht aufrechterhalten lässt.

8 Entwicklung von Messinstrumenten zu Inklusion

Im Folgenden geht es darum zu verstehen, welche Rahmenbedingungen (→ Abb. 8.1 die Zahnräder) die Deutung von Inklusion und die Entwicklung von Instrumenten (→ Abb. 8.1 der rechte Kasten) geprägt haben. Dabei liegt der Fokus auf dem individuellen Sensemaking der Instrumenteentwickler_innen. Gleichwohl gibt es Hinweise darauf, wie die Aushandlungsprozesse auf kollektiver Ebene abgelaufen sind und welche Aspekte hierbei eine Rolle gespielt haben.

Beim Sensemaking in Organisationen finden mehrere Prozesse zeitgleich statt (nähere Ausführung → Kapitel 5):

- **Auf individueller Ebene: Wenn ein einzelner Sensemaker etwas deutet**, greift er auf verschiedene Rahmenbedingungen aus seinem Referenzrahmen zurück, z. B. vorhandene Wissensbestände, persönliche Erfahrungen oder relevante Rahmenbedingungen. Die jeweiligen Rahmenbedingungen interagieren im Deutungsprozess miteinander. Die einzelnen Sensemaker können sich darin unterscheiden, welche Wissensbestände (→ Abb. 8.1 die Zahnräder) sie bei der Deutung der gleichen Sache – in diesem Falle Inklusion – für relevant erachten. Demnach variieren die Rahmenbedingungen zwischen den Sensemakern. Dennoch gibt es häufig auch ähnliche Wissensbestände, die genutzt werden, um etwas zu deuten.
- **Auf Bundeslandebene:** Neben dem Sensemaking auf individueller Ebene laufen in Organisationen zeitgleich noch kollektive Aushandlungsprozesse **auf der intersubjektiven Ebene ab**. Dies wäre im Fall von Schulinspektionen der Moment, wenn mehrere Personen zusammenkommen, um gemeinsam Instrumente für Inklusion zu entwickeln. Dabei ist zu berücksichtigen, dass der kollektive Aushandlungsprozess das Sensemaking von Individuen mit definiert und vice versa.

All diese Prozesse, die im Rahmen des Sensemakings in Organisationen stattfinden, können als zirkulär beschrieben werden. Die schriftliche Ausformulierung, die nun erfolgt, erfordert es jedoch, dass die Prozesse sequentiell beschrieben werden. Der zirkuläre Charakter der Prozesse wird hierbei berücksichtigt, indem immer wieder Verweise auf zeitgleich stattfindende Prozesse oder prägende Rahmenbedingungen aufgeführt werden. In den einzelnen Teilkapiteln wird der Gegenstand der Arbeit aus verschiedenen Perspektiven betrachtet.

So wird **im ersten Schritt** (→ Kapitel 8.1) dargestellt, welche Rahmenbedingungen in Bezug auf das Sensemaking von Inklusion von den Interviewpartnern genannt wurden. Dazu wird auf die drei Dimensionen zurückgegriffen, die nach Spillane et al. (2002) das Sensemaking von politischen Vorgaben prägen: „persönliche Attribute des Sensemakers“, der „situative Kontext“ sowie „politische Vorgaben“. Ausgehend von der Abbildung 8.1 geht es um eine Beschreibung dessen, **was sich bundesländerübergreifend und über den Individualfall hinausgehend hinter den „Zahnrädern“ verbirgt**. Hierfür wird aufgezählt, welche konkreten Rahmenbedingungen bei der Entwicklung von Instrumenten zu Inklusion eine Rolle gespielt haben. Zugleich wird dargestellt, inwiefern die Rahmenbedingungen in ihren Impulsen im Hinblick auf die Deutung von Inklusion zwischen den Einzelinterviews variieren. Um ein Beispiel zu nennen: Bei fast allen Interviewpartnern wurde das Kultusministerium als eine Rahmenbedingung genannt, welche das Sensemaking zu Inklusion bzw. die Bewertungsinstrumente zu Inklusion mitdefinierte.

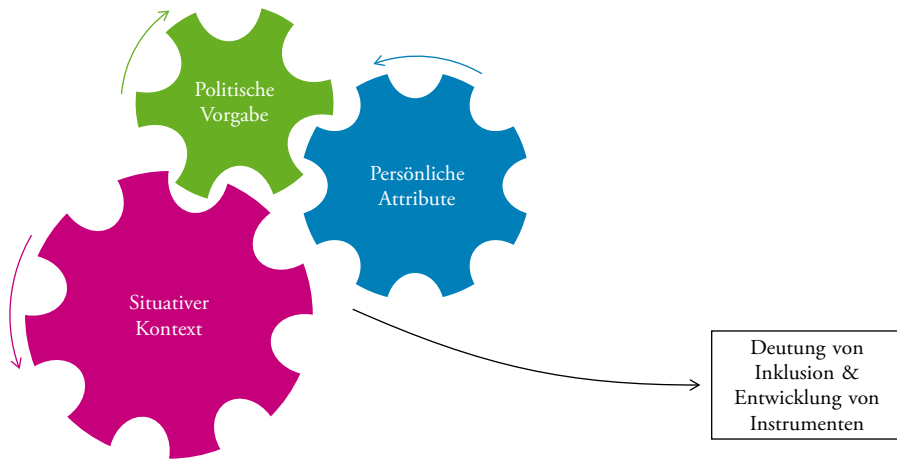


Abb. 8.1: Sensemaking von Inklusion

Jedoch gibt es Unterschiede darin, wie diese das Setting prägen, z. B. unterscheiden sich die Kultusministerien darin, welche Hinweise und Vorschläge sie in Bezug auf Inklusion gemacht haben (-Impulse) bzw. inwiefern diese von den Instrumenteentwickler_innen berücksichtigt wurden.

Zusätzlich wird in diesem Block auch dargestellt, welche Rahmenbedingungen die Entwicklung von Instrumenten im Allgemeinen prägen. Dies dient als Hintergrundinformation, da es bislang kaum Forschung zur Entwicklung von Messinstrumenten gibt. Darüber hinaus ermöglicht dies erklären zu können, wie die Instrumente zustande gekommen sind, die zwar nicht explizit als „inklusionsorientiert“ gekennzeichnet sind, aber Teilaspekte von Inklusion repräsentieren (→ Abschnitt 7.1). Des Weiteren betreffen viele Rahmenbedingungen, die für das gesamte Instrumentenset gelten, auch die Instrumente von Inklusion, z. B. ob es generell einen Wahlbereich im Instrumentenset gibt, hat auch Implikationen in Bezug auf die Instrumente zu Inklusion.

Im zweiten Schritt (→ Abschnitt 8.2) wird eine neue Perspektive eingenommen: Ausgehend von einzelnen Teilaspekten von Inklusion (→ Abb. 8.1 der rechte Kasten), z.B. der Zielgruppe, geht es nun um die Frage, welche Rahmenbedingungen (→ Abb. 8.1 die Zahnräder) in Bezug auf diese Teilaspekte prägend wirkten.

Im dritten Schritt (→ Abschnitt 8.3) wird die Interaktion zwischen verschiedenen Rahmenbedingungen bei einzelnen Instrumenteentwickler_innen beschrieben. Demnach liegt nun der Fokus auf individuellen Sensemakern. Ausgehend von den in Kapitel 7.3 dargestellten Typen von Inklusionsdeutungen wird exemplarisch das Sensemaking von vier Personen beschrieben. Indem die Interaktion von Rahmenbedingungen bei einzelnen Personen beschrieben wird, wird u. a. nachvollziehbar, inwiefern unterschiedliche Rahmenbedingungen zu ähnlichen Deutungen von Inklusion führen können. Jedoch kann es auch sein, dass ähnliche Rahmenbedingungen zu unterschiedlichen Deutungen führen.

Der vierte Schritt (→ Abschnitt 8.4) der Auswertung bewegt sich nicht mehr auf der Ebene der einzelnen Instrumenteentwickler_innen, sondern auf der Bundeslandebene. Hier stehen nun die kollektiven Aushandlungsprozesse in den einzelnen Bundesländern im Mittelpunkt. Diese

Analyse leistet einen Beitrag zu der Frage, warum sich bestimmte Deutungen der einzelnen Instrumenteentwickler_innen in den Instrumenten der bundesländerspezifischen Schulinspektionen durchsetzen bzw. nicht durchsetzen konnten.

Im fünften Schritt (→ Abschnitt 8.5) wird im Rahmen eines Zwischenfazit dargestellt, welche Handlungsmuster sich in Bezug auf die Entwicklung von Instrumenten bundesländerübergreifend und über den Individualfall hinausgehend herausbilden lassen.

Die Entwicklung von Bewertungsinstrumenten in den einzelnen Schulinspektionen ähneln sich: Meistens gibt es eine Arbeitsgruppe von Personen, die für die konkrete Formulierung von Bewertungsinstrumenten zuständig ist bzw. die Koordination von kleineren Arbeitsgruppen übernimmt. Hierbei sind neben den Instrumenteentwickler_innen der Schulinspektionen unterschiedliche Akteure beteiligt, z. B. Vertreter der Ministerien, der Schularten, die Schulaufsicht, Forscher_innen usw. Zusätzlich finden eine Sichtung der relevanten Literatur sowie ein Erfahrungsaustausch mit anderen Bundesländern bzw. anderen Ländern außerhalb der Bundesrepublik Deutschland statt. Während des Prozesses erhalten die Arbeitsgruppen Rückmeldungen von verschiedenen Akteuren, passen daraufhin die Instrumente an und spiegeln die überarbeiteten Versionen wieder zurück. Es gibt jedoch zwischen den Bundesländern Unterschiede in Bezug darauf, wer als Akteur einbezogen wird und wer zu welchem Zeitpunkt des Prozesses mitentscheiden darf.

8.1 Prägende Rahmenbedingungen beim Sensemaking

Im Rahmen der Arbeit geht es um die Frage, welche Rahmenbedingungen das Sensemaking von Inklusion geprägt haben (für eine Übersicht → Tabelle 8.1). In Bezug auf manche Rahmenbedingungen ergaben sich bereits vor der empirischen Analyse aus den theoretischen Vorüberlegungen, z. B. Schulaufsicht als Akteur, Hinweise. Dagegen kristallisierten sich andere Rahmenbedingungen, z. B. „schulische Vertreter“ während der Analyse als relevante Kategorien heraus. Bestimmte Rahmenbedingungen haben in fast allen Interviews eine Rolle gespielt, z. B. das Kultusministerium, und andere Rahmenbedingungen wirkten eher selten prägend, z. B. „Behindertenbeauftragte“. Im Folgenden wird auf die einzelnen Rahmenbedingungen näher eingegangen.

8.1.1 Persönliche Attribute der Entwickler_innen

Bei der Entscheidung darüber, welche persönlichen Attribute von Relevanz sind, war maßgeblich, ob in den Daten ein Zusammenhang zwischen Deutung von Inklusion und den jeweiligen Attributen herstellbar war. Im Folgenden werden als relevante Rahmenbedingungen „Qualifikation und Erfahrungen aus der Praxis“, „Erfahrungen als Inspektor_in“ sowie „Verständnisse von guter Schule“ näher beschrieben. Diese wirkten sich darauf aus, wie Inklusion definiert wird und wie sich die Instrumenteentwickler_innen zur Umsetzung von Inklusion positionieren.

Qualifikation und Erfahrungen aus der Praxis: In den Interviews zeigte sich, dass Expertise im sonderpädagogischen Bereich für viele Interviewpersonen für die Deutung von Inklusion relevant war: So gab ein Drittel der Interviewpartner an, dass sie von Hause aus Sonderpädagog_innen sind (Int. 3; 7; 14; 15;) bzw. eine sonderpädagogische Zusatzqualifikation im Rahmen der Erstausbildung und nach dem Studium (Int. 9) erhalten haben. Des Weiteren betonten einige Interviewpersonen ihre Praxiserfahrung im gemeinsamen Unterricht von Schü-

Tab. 8.1: Rahmenbedingungen beim Sensemaking

Dimensionen nach Spillane et al.		Konkrete Rahmenbedingungen beim Sensemaking zu Inklusion
Persönliche Attribute des Sensemakers		<ul style="list-style-type: none"> • Qualifikation und Erfahrungen aus der eigenen Lehrpraxis • Erfahrungen als Inspektor_in • Verständnis von guter Schule
Rechtliche Vorgaben als Form von politischen Vorgaben		<ul style="list-style-type: none"> • BRK • Weitere rechtliche Vorgaben
Situativer Kontext	Akteure	<ul style="list-style-type: none"> • Kultusministerium • Schulische Vertreter • Schulaufsicht • Schulinspektionen in anderen Bundesländern • Forscher_innen • Schüler_innen und Eltern • Weitere Akteure
	Kontextbedingungen	<ul style="list-style-type: none"> • Öffentliche Debatte • Umsetzung von Inklusion im jeweiligen Bundesland • Schulqualitätsrahmen • Bewertungsinstrumente aus dem ersten Zyklus • Zeitliche sowie finanzielle Bedingungen • Funktionen von Schulinspektionen • Empirische Messung von Schulqualität

ler_innen mit und ohne sonderpädagogischen Förderbedarf (z. B. Int. 9_14; 3_7; 15_1). So antwortet eine Instrumenteentwicklerin auf die Frage, woher sie Input erhalten hat: „*Ich merke gerade - ich war wirklich ganz stark in diesem Inklusionsthema drin und habe da eben vom Index zu Inklusion ganz viele Informationen übernommen und (...) natürlich auch nen Teil aus dem aktiven Schulalltag, weil ich noch einige Tage an der Schule bin*“ (Int. 3_7). Dabei waren die jeweiligen Personen in der Praxis häufig nicht nur als Lehrkraft, sondern auch im Schulleitungsteam tätig. Demnach verfügten sie auch über persönliche Erfahrungen im Bereich Schulmanagement. Bei der Frage danach, welche Themen durch die Instrumente zu Inklusion abgedeckt werden sollen, griffen sie häufig auf ihre Erfahrungen aus der Praxis zurück: „*Wir haben andere Kollegen, die eben auch Erfahrung im inklusiven Setting haben und wir haben einfach das, was es so an Praxiserfahrung gibt, zusammengenommen und ähm – im Grunde – aufgrund dieser realen Anlässe, wo Inklusion sichtbar wird an Schulen, das definiert*“ (Int. 2_21).

Die Relevanz von sonderpädagogischer Expertise zeigt sich darüber hinaus dadurch, dass denjenigen Personen, die darüber verfügen, innerhalb der Inspektion meistens die Zuständigkeit für die Themenbereiche Inklusion sowie Förderschule zugeschrieben wurde (z. B. Int. 15_1; 3_1; 8_21; 9): „*Also ich bin die Sonderschullehrerin im Team [..., ap] Und da war klar, dass die ganzen*

sonderpädagogischen Themen, aber auch natürlich das große Thema Inklusion – das ja eigentlich kein sonderpädagogisches Thema ist (.), aber Inklusion steht so quasi – das ist mein Eindruck – an meiner Zimmertür“ (Int. 3_1). Es gab auch den Fall, dass Personen mit sonderpädagogischer Qualifikation nach der Ratifizierung der BRK eingestellt wurden, damit dieses Thema von Seiten der Schulinspektion abgedeckt werden kann (z. B. Int. 9). Auch wenn diesem Personenkreis nicht die alleinige Zuständigkeit für Inklusion obliegt, wird den sonderpädagogischen Fachkräften im Team die Expertise bei der Entwicklung von Instrumenten zu Inklusion zugeschrieben: „Es gibt hier diese Fachgruppe der – der Sonderpädagogen, die ham vorbereitet. Beziehungsweise, damals gab s auch ähm (.) eine Arbeitsgruppe Inklusion, die also besetzt war mit Förderschulmensen und Regelschulleuten. Und wir ham – wir hams vorbereitet, dann im Kollegium vorgestellt, abgestimmt“ (Int. 15_7).

Die Relevanz von sonderpädagogischer Expertise hat in manchen Schulinspektionen auch zur Folge, dass sich nur ein Teil der Instrumenteentwickler_innen überhaupt mit Inklusion auseinandergesetzt hat: „Also ich kenn mich bei Inklusion überhaupt nicht aus, ich kann nur sagen, die Experten, die sich damit befassen, haben mir dieses Merkmal zugeliefert und im Grunde die Qualitätserwartungen formuliert“ (Int. 8_21, siehe auch 18_26). Während in diesem Fall auch kein Bedarf an weiterer Qualifizierung in Bezug auf Inklusion geäußert wird, sehen andere Instrumenteentwickler_innen noch Qualifizierungsbedarf bei sich selbst oder anderen im Team: „Auf jeden Fall wird immer wieder geguckt, welche Expertise brauchen eigentlich alle, was diese Basis ist. Wir sind nicht automatisch alle Förderschullehrkräfte und können einschätzen, was braucht ein einzelnes Kind. Aber wir brauchen ein Gespür dafür, inwieweit eine Auseinandersetzung mit den Bedarfen einzelner Kinder an der Schule angemessen wäre. Ob jetzt genau diese Hörhilfe oder Sehunterstützung oder Mobilitätsunterstützung genau für den Bedarf an diesem Kind passt, das ist schwierig, wenn man nicht den Hintergrund hat. Aber wir arbeiten seit längerer Zeit daran allen Inspektor_innen [anonymisiert] so ein Basiswissen zur Verfügung zu stellen“ (Int. 5_31; siehe auch 8_21; 5_31; 9_12; 8_18; 18; 10_39; 7_24). In dem Zusammenhang haben die Mitarbeiter_innen der Schulinspektionen auch im Blick, dass Expertise nicht nur für die Entwicklung, sondern auch für die Anwendung der Instrumente notwendig ist. So wird auch als Begründung für die Auswahl von Inhalten genannt, dass Inspektor_innen eine bestimmte Expertise benötigen (Int. 7_28; siehe auch 9_48).

Zusammenfassend zeigt sich, dass es häufig Personen mit sonderpädagogischer Qualifizierung und mit Erfahrung im gemeinsamen Unterricht sind, die für den Themenbereich „Inklusion“ zuständig sind, bzw. denen Expertise zugeschrieben wird. Die Anerkennung von sonderpädagogischen Fachkräften als Expert_innen in Bezug auf Inklusion legt nahe, dass viele unter Inklusion den gemeinsamen Unterricht von Schüler_innen mit und ohne Behinderung verstehen (→ Typ I und Typ II in Abschnitt 7.3). Es gibt aber auch Instrumenteentwickler_innen mit sonderpädagogischer Expertise, die sich von dieser Deutung von Inklusion distanzieren (nähere Ausführung → Abschnitt 8.3). Demnach ergibt sich aus der sonderpädagogischen Qualifizierung und der Erfahrung im gemeinsamen Unterricht nicht automatisch ein enger Inklusionsbegriff. Des Weiteren zeigt sich, dass die Erfahrungen aus der Praxis für die inhaltliche Gestaltung der Instrumente von Relevanz sind. Dies spiegelte sich auch bei Hospitationen der Autorin dieser Arbeit wider: So war sie dabei, als ein Brainstorming über neue Instrumente zur Integration von Menschen mit Fluchterfahrung stattfand. Dabei wurde deutlich, dass auch in diesem Fall die Erfahrungen aus der Praxis eine bedeutende Bezugsquelle waren, um neue Instrumente zu formulieren.

Erfahrung als Inspektor_in: Innerhalb der Schulinspektionen gibt es unterschiedliche Zuständigkeitsbereiche: Während die Inspektor_innen primär dafür zuständig sind, die Schulen zu besuchen und die Evaluationen durchzuführen, sind die wissenschaftlichen Referent_innen unter anderem für die Einhaltung von wissenschaftlichen Gütekriterien zuständig. Beide Gruppen sind an der Entwicklung von Instrumenten beteiligt, jedoch ist zu berücksichtigen, dass nicht in allen Inspektionen wissenschaftliche Referent_innen tätig sind. Die unterschiedlichen Zuständigkeiten führen auch zu unterschiedlichen Perspektiven im Hinblick auf die Entwicklung von Instrumenten: *„Wenn man das sozusagen rein empirisch mit der Brille macht und rein mit der praktischen Brille. Das sind einfach zwei verschiedene Brillen. Die eine ist durch und durch praxisgesättigt, die andere ist durch und durch empirisch-theoretisch gesättigt, da kommt man zu unterschiedlichen Entschlüssen“* (Int. 13_21).

Im Folgenden liegt der Fokus zunächst auf der Perspektive der Inspektor_innen. Innerhalb der Schulinspektion wird den Inspektor_innen als Expertise zugeschrieben, dass sie beurteilen können, inwiefern die Bewertungsinstrumente in der Praxis anwendbar sind, z. B. *„Ähm ist dann an die Inspektoren gegangen und auch da gab es mehrere Schleifen und an der Stelle ähm tut sich dann auch wirklich immer Arbeit auf, weil der wissenschaftliche Blick ist einer, aber der Menschen, die dann wirklich Expertise haben und auch in der Praxis agieren müssen und gucken müssen, können wir wirklich mit diesen, ich nenne es jetzt Kriterien, gut arbeiten, ist nochmal was anderes“* (Int. 16_10).

So bringen diese bei der Entwicklung von Instrumenten – wie bereits oben dargestellt – ihre Erfahrungen aus der eigenen Praxis als Lehrkraft ein sowie im Rahmen ihrer Tätigkeit als Inspektor_innen ihr Erfahrungswissen darüber ein, was die Schulen im jeweiligen Bundesland aktuell bewegt (Int. 2_21; 10_25; 4_16). Demnach verfügen die Inspektor_innen – im Gegensatz zu den wissenschaftlichen Referent_innen – über Expertise darüber, welche Inhalte sinnvoll sind, um den Schulen Impulse für Schulentwicklung geben zu können (Int. 4_10; 10_2; 7_8; 13_21). Des Weiteren bringen sie bei der Entwicklung von Instrumenten ihre Einschätzung darüber ein, welche sozialen Phänomene im Alltag beobachtbar und trennscharf zu erheben sind (Int. 13_21) (nähere Ausführung → Abschnitt 8.1.3). Jedoch gehen mit dieser Perspektive auch Grenzen einher: So gibt es bei den Inspektor_innen und den wissenschaftlichen Referent_innen unterschiedliche Positionen zu der Frage, ob bestimmte Themen bzw. deren Operationalisierung in der Praxis erfasst werden können: *„Noch n Anlauf wieder genommen, also es war auch dann schwierig, weil die Fragen teilweise nicht den wissenschaftlichen Standards entsprachen, die die Wissenschaftler haben mit den Fragen, die die Praktiker eher im Interview fragen würden. Wir haben dann mehrere Schleifen gehabt, immer wieder versucht das zusammen zu bringen. Die Wissenschaftler haben gesagt, das geht so nicht, die Praktiker haben gesagt so können wir es nicht fragen“* (Int. 16_33). Demnach stimmen die Einschätzungen der Inspektor_innen zum Teil nicht damit überein, welche Anforderungen von Seiten der wissenschaftlichen Referent_innen an die Instrumente formuliert werden.

Verständnisse von guter Schule: Auf die Frage, was die Instrumenteentwickler_innen unter einer „guten“ bzw. einer „schlechten“ Schule verstehen, gibt es verschiedene Reaktionen: Bei vielen Interviewpersonen zeigt sich als erste Tendenz, dass sie einer konkreten Positionierung ausweichen möchten. Auf die Frage, was eine gute Schule sei, antwortet beispielsweise eine Interviewperson: *„Ähm, (.) das is – das is glaube ich wirklich schwierig so einfach zu beschreiben. Also ich denke, gute Schule muss – äh muss für mich was – was Lebendiges sein“* (Int. 15_8). Viele Interviewten nennen als erstes, dass eine Schule gut sei, wenn der Unterricht gut sei

(Int. 8; 15; 18; 4_13; 6; 7 10_22; 18_29), z. B. *„Und das Kerngeschäft, wie wir auch oft sagen, das Kerngeschäft einer jeden Schule ist der Unterricht, ist guter Unterricht“* (Int. 6). Im weiteren Verlauf des Interviews treffen viele Interviewpartner konkretere Aussagen, die die Qualität von Schule auf der Prozessebene bzw. den anzustrebenden Output von Schule beschreiben. Dabei lassen sich auf der Prozessebene zwei Merkmale anführen, die von mehreren Personen genannt werden:

Erstens ist dies der Umgang mit Heterogenität. So wird es als guter Unterricht gesehen, wenn auf die Heterogenität der Schüler_innen Rücksicht genommen und ein differenziertes Bildungsangebot gemacht wird (Int. 6, 16; 15_3; 12; 7). *„Im Blick auf Unterricht hoffe ich, dass ähm zunehmend der Blick dahingeht, dass der Unterricht so ausgerichtet sein muss, dass er der Heterogenität und ich sage bewusst nicht Inklusion, ähm der Schüler gerecht wird“* (Int. 16). Als Begründung dafür, warum der Umgang mit Heterogenität sinnvoll ist, wird *„allen Teilhabe ermöglichen“* (Int. 15) sowie *„Chancengerechtigkeit“* (Int. 12) genannt.

Als zweite Beschreibung von guter Schule wird ein Verständnis von Schule als lernende Organisation genannt. Dabei wird der Anspruch an die Schule formuliert, dass sie regelmäßig die eigenen Prozesse reflektiert, diese in Abhängigkeit von den gesetzten Zielen evaluiert und davon ausgehend Prozesse optimiert (Int. 10_13; 8; 12; 2; 13; 1_46; 6; 14_15): *„Gute Schule solche, die ne gut funktionierende Evaluationskultur haben oder Reflexionskultur haben. Die sich also selbst gut steuern können und das erfolgreich tun“* (Int. 10). Diese Beschreibung von Qualität lässt sich theoretisch auch auf andere Organisationen anwenden, z. B. Unternehmen oder Verwaltung. So bleibt in Bezug auf die ethisch-moralische Dimension von Bildungseinrichtungen eine Leerstelle. Es fehlt an einer Setzung, welche Ziele mit den jeweiligen Lernprozessen von Schule erreicht werden sollen. In diesem Zusammenhang kritisiert Biesta:

„The problem is, however, that effectiveness is an instrumental value, a value which says something about the quality of processes and, more specifically, about their ability to bring about certain outcomes in a secure way. But whether the outcomes themselves are desirable is an entirely different matter – a matter for which we need value-based judgements that are not informed by instrumental values but by what we might best call ultimate values“ (Biesta 2009, S. 35).

In den Interviews zeigt sich, dass diese Leerstelle nur zum Teil mit Inhalten gefüllt wird. Als weitere Merkmale auf der Prozessebene werden noch die Kooperation von Lehrkräften (Int. 1_46), die Unterrichtskultur (Int. 1_46) sowie die Zufriedenheit mit Schule (Int. 12_19) genannt. Des Weiteren wird eine Vielzahl von Zielen von guter Schule formuliert: Chancengerechtigkeit (Int. 12_19), Erziehung zu mündigen Bürger_innen (Int. 9_10 ; 16_21; 5_13; 7_9), Lehren und Lernen (Int. 18_9; 12_20); Wohlfühlen (Int. 1_46; 12_20), berufliche Orientierung (Int. 12_20) sowie Qualifizierung zum guten Leben (Int. 10_14): *„Also sie [die Schule, ap] muss gucken, was wollen wir denn, also warum sind wir überhaupt eigentlich da. Also im weitesten Sinne natürlich, um Menschen zu befähigen gutes Leben zu führen, nicht? Sich selbst versorgen zu können“* (Int. 10).

In den Interviews werden unterschiedliche Beschreibungen von guter Schule genannt, die jedoch auf einer abstrakten Ebene bleiben. Demnach könnten die einzelnen Beschreibungen eher als Aufzählung verstanden werden, da sie sich auf den ersten Blick eher ergänzen als dass sie sich widersprechen. Erste Widersprüche wurden jedoch an anderen Stellen der Interviews sichtbar: So gab es unterschiedliche Positionen in Bezug auf die Frage, ob die optimale Förderung durch Selektion (Int. 8_30) oder durch gelungenen Umgang mit heterogenen Gruppen (Int. 15_9) gelingt.

Im Folgenden geht es nun um die Frage, inwiefern das eigene Verständnis von guter Schule der Interviewpartner Überschneidungen mit dem in dieser Arbeit verwendeten Verständnis von Inklusion aufweist. In diesem Zusammenhang konnten drei unterschiedliche Muster herausgearbeitet werden:

- So zeigt sich bei manchen Interviewpartnern die eindeutige Positionierung, dass Inklusion mit dem eigenen Verständnis von guter Schule nicht kompatibel sei. Dies sind häufig Vertreter des Typ II „Inklusion als Herausforderung für Schulen“ (nähere Ausführung → Abschnitt 7.3) und lässt sich dadurch erklären, dass die Personen ein selektives Schulsystem favorisieren: *„Ich bin also, wenn Sie so wollen, Anhänger gymnasialer Bildung. Und das ist kein inklusiver Ansatz“* (Int. 8_30). In diesen Fällen hat eine Auseinandersetzung zu der Frage stattgefunden, inwiefern Inklusion zu dem eigenen Verständnis von guter Schule passt und das Ergebnis ist eine kritische Positionierung gegenüber Inklusion. Dies spiegelt sich auch in der Position dazu wider, welche Relevanz Instrumenten zu Inklusion in dem kompletten Instrumentenset der Schulinspektionen zukommen sollte.
- Als weiteres Muster zeigt sich, dass Inklusion als etwas verstanden wird, das Überschneidungen zu dem eigenen Verständnis von guter Schule aufweist. Dieses Muster ist insbesondere bei Typ III „Inklusion als gelungener Umgang mit Heterogenität“ (nähere Ausführung → Abschnitt 7.3) vorzufinden. Demnach hat sich aus unterschiedlichen Gründen ein bestimmtes Verständnis von guter Schule herausgebildet und Inklusion wird so gedeutet, dass es zu diesem Verständnis passt. Somit wird von Seiten der Instrumenteentwickler_innen auch eine positive Haltung gegenüber Inklusion eingenommen.
- Bei manchen Interviewpartnern wurde eine direkte Verknüpfung zwischen ihrem Verständnis von guter Schule und Inklusion in der Analyse nicht sichtbar. Eine mögliche Erklärung könnte hierfür sein, dass Inklusion nicht als ein Indikator für gute Schule gesehen wird. Hierbei entsteht der Eindruck, dass die Auseinandersetzung mit Inklusion unabhängig davon stattfand, was unter einer guten Schule verstanden wird. Dies legt nahe, dass es wahrscheinlich keine Widersprüche zum vorhandenen Verständnis von guter Schule gab, aber auch keine intensive Auseinandersetzung über die normative Aussagekraft von Inklusion in Bezug auf Schulqualität stattgefunden hat. Dies ist insofern überraschend, als die Instrumente der Schulinspektion zur Messung von Qualität genutzt werden. Ausgehend von dem Sensemaking-Ansatz wäre es theoretisch auch möglich, dass sich durch die Auseinandersetzung mit Inklusion das eigene Verständnis von guter Schule verändert hat. Hierzu gibt es jedoch in den Interviews keine Hinweise. Des Weiteren wird auch im Rahmen des Sensemaking-Ansatzes argumentiert, dass es selten zu Veränderungen in Bezug auf bestehende Überzeugungen kommt (nähere Ausführung → Abschnitt 9.2).

8.1.2 Rechtliche Vorgaben

Neben den persönlichen Attributen spielt beim Sensemaking auch eine Rolle, wie die politischen Vorgaben bzw. deren Repräsentation rezipiert wurden. Hierzu wird im Folgenden auf die BRK, die maßgeblich die Debatte zu Inklusion im deutschsprachigen Raum initiiert hat, eingegangen sowie auf weitere rechtliche Vorgaben wie die Schulgesetzgebung in den einzelnen Bundesländern.

UN-Behindertenrechtskonvention: Die Ratifizierung der BRK hat bei den Schulinspektionen verschiedene Prozesse initiiert, die im Folgenden näher dargestellt werden: So hat diese

dazu geführt, dass eine Auseinandersetzung über Bewertungsinstrumente zu Inklusion stattgefunden hat (z. B. Int. 2_27; 3_24; 9_47; 8_24; 15_2; 11_25): *„Ja in der ersten Runde (.) ähm, ham wir eigentlich das Thema Inklusion (.) überhaupt nich im Blick gehabt. In der zweiten Runde ham wir dann zum einen, äh, noch mal in einer längeren Phase tatsächlich überprüft, ob wir mit den vorhandenen Bewertungskriterien Inklusion abbilden können“* (Int. 15_2). Die Auseinandersetzung mit Inklusion, die häufig mit dem Beginn des zweiten Inspektionszyklus zusammenfiel, lässt sich auch darauf zurückführen, dass die BRK zu Veränderungen in der schulischen Praxis sowie in der Schulgesetzgebung geführt hat und sich somit indirekt auf das Handeln der Schulinspektion auswirkte (Int. 5_36; 6_18; 2_27). So antwortet eine Person auf die Frage, welche Rolle die BRK gespielt hat: *„Für unsere Arbeit jetzt speziell, kann ich nicht sagen. Also das ist eher die Behindertenrechtskonvention, äh hat ja letztlich die gesetzlichen Regelungen hier in [Bundesland] sehr stark geprägt. Also äh die Tatsache, dass inklusive Bildung so eine entscheidende Rolle spielt in unseren Schulen geht ja zurück auf die UN-Konvention“* (Int. 6_18).

In diesem Zitat wird bereits angedeutet, was sich als Muster in vielen Interviews herauskristallisiert: So zeigt sich, dass die konkreten Inhalte der BRK bei der Entwicklung von Instrumenten eine geringe Rolle gespielt haben: *„Sie ist ja im Grunde die Grundlage für das Schulgesetz an der Stelle und das ist natürlich für uns auch wieder bindend, aber spezielle Teile dort. Also es ist jetzt nicht so, dass wir das jetzt beiseitegelegt haben und gesagt haben, wie erfüllen wir die und uns dann eher so, wie es umgesetzt wird an den Schulen, worauf wir gucken“* (Int. 1_38; siehe auch 2_27; 9_47; 12_29). Ergänzend bestätigen einige Befragte, dass die BRK insgesamt keinen Einfluss auf ihre Arbeit gehabt hätte (Int. 13_26; 4_12; 16_53) und somit auch nicht auf inhaltlicher Ebene. So gibt es lediglich zwei Fälle, die davon sprechen, dass man sich mit den konkreten Inhalten von Art. 24 BRK auseinandergesetzt habe (Int. 14_35; 1_37), aber auch hier gibt es kaum Hinweise darauf, welche Implikationen dies in Bezug auf die konkrete Formulierung von Indikatoren hatte.

Generell zeichnet sich demnach als Muster ab, dass die Inhalte der BRK für das Entwickeln von konkreten Instrumenten kaum eine Rolle gespielt haben, aber die Ratifizierung der BRK bei der Schulinspektion Prozesse initiiert hat (Int. 2_27; 9_47; 8; 15_2; 11_25; 3_24). In diesem Zusammenhang stellt sich die Frage, wie sich der geringe Einfluss der BRK auf inhaltlicher Ebene erklären lässt. Hierzu gibt es in den Interviews insofern Hinweise, als eine Person dies damit begründet, dass die BRK zu weit vom Schulalltag entfernt sei (Int. 2_27), d.h. dass sich die Inhalte nur schwer direkt auf den Alltag in der Schule übertragen lassen. Als weitere Begründung wird genannt, dass keine Notwendigkeit bestanden habe sich mit der BRK auseinanderzusetzen, da die Schulinspektion sich selbst nicht als Kontrollinstanz und somit nicht in der Verantwortung sehe, das Monitoring von Art. 24 BRK zu realisieren (Int. 5_36). So sei dies die Aufgabe der Schulaufsicht (Int. 12_12; nähere Ausführung zur Schulaufsicht → Kapitel 8.1.3). Jedoch gibt es in den Interviews auch Positionierungen, die in eine andere Richtung gehen: Diese können Typ 1 „Inklusion als Anspruch auf Förderung“ (nähere Ausführung → Kapitel 7.3) zugeordnet werden. Jedoch gibt es auch bei diesen Interviewpersonen keinen Hinweis darauf, dass die BRK für die inhaltliche Gestaltung der Instrumente genutzt wurde.

Andere rechtliche Vorgaben: Die BRK stellt eine internationale Konvention dar und es fehlen Angaben darüber, welche konkreten Implikationen sich aus den rechtlichen Vorgaben der BRK für Deutschland ergeben (vgl. Piezunka et al. 2018). Möglicherweise waren im Hinblick auf die Entwicklung der Instrumente rechtliche Vorgaben auf nationaler bzw. bundesland-spezifischer Ebene von Bedeutung. Jedoch berichten die Instrumenteentwickler_innen, dass

auch die Inhalte aus Schulgesetzen und anderen Erlassen bei der inhaltlichen Entwicklung von Instrumenten eine geringe Rolle gespielt hätten (Int. 2_25; 3_10; 1_21; 4_12; 9_22; 14_8). Dies wird in Bezug auf Inklusion in einem Bundesland damit begründet, dass sie zu einem Zeitpunkt vor der Herausforderung standen, Bewertungsinstrumente zu entwickeln, als die rechtlichen Verordnungen noch nicht verabschiedet waren. Des Weiteren begründen andere Instrumenteentwickler_innen den fehlenden Einfluss der rechtlichen Vorgaben erneut damit, dass sie die Schulinspektion nicht als einen Akteur ansehen, der das Monitoring von Gesetzen übernimmt (Int. 9_22; 4_8; 3_26; 12_16).

Auch wenn sich aus den rechtlichen Vorgaben für die meisten Interviewpartner keine Inhalte für die Bewertungsinstrumente ergeben, ist der Anspruch vorhanden, dass die jeweiligen Messinstrumente mit dem rechtlichen Rahmen konform sind (Int. 14_8; 12_17; 9_20; 5_9; 1_43): „Ja gut - wir sind natürlich wirklich in diesem Rechtsrahmen unterwegs und wir müssen natürlich auch in diesem Rahmen bleiben“ (Int. 1_43). So wird häufig nachträglich auf die rechtlichen Rahmenbedingungen verwiesen, um einzelne Instrumente zu legitimieren (nähere Ausführung → Abschnitt 9.2) „Das [-Hospitanz durch die Schulleitung, ap] sind gesetzliche Vorgaben, die wir dann auch verankert haben. Also es gibt schon ne Reihe von Bewertungsinstrumenten [anonymisiert, ap] - aber nicht alle - die auch mit Erlassen oder Gesetzen unterlegt sind, die wir dann auch zur Grundlage gemacht haben. Sachen, wenn das nicht passiert, dann kann der Prozess nicht funktionieren“ (Int. 10_12, siehe auch 11_11).

Zusammenfassend wird deutlich, dass Gesetzesveränderungen zwar dazu führen, dass Prozesse initiiert werden, z. B. Auseinandersetzung mit Inklusion oder Einstellung von neuem Personal. Wenn es jedoch um die Inhalte von Instrumenten geht, spielen die konkreten Gesetzestexte eine geringe Rolle. In manchen Fällen dienen sie retrospektiv zur Legitimation von einzelnen Anforderungen (Int. 15_28; 1_17) oder zur Festlegung der Höhe der Anforderungen (Int. 15_28; 5_38; 16_37), die von Seiten der Schulinspektion gegenüber den Schulen formuliert werden.

8.1.3 Situativer Kontext

Neben den Attributen der Sensemaker sowie den rechtlichen Vorgaben, prägt auch der situative Kontext das Sensemaking in Bezug auf Inklusion. Spillane et al. (2002) lassen relativ offen, welche Aspekte unter den situativen Kontext fallen, und nennen eine Vielzahl von unterschiedlichen Rahmenbedingungen, die eine Rolle spielen können. In Bezug auf die Deutung von Inklusion bzw. die Entwicklung von Messinstrumenten zu Inklusion zeigte sich in den Interviews, dass folgende Akteure von besonderer Relevanz sind: Kultusministerien, Schulen, Forscher_innen; Schulinspektionen aus anderen Ländern sowie die Schulaufsicht. Als relevante Kontextbedingungen sind unter anderem die öffentliche Debatte zu Inklusion, die Umsetzung von Inklusion im Bundesland sowie zeitliche und finanzielle Bedingungen zu nennen. Im Folgenden wird zunächst dargestellt werden, welche Akteure bei der Entwicklung von Instrumenten miteinbezogen bzw. welche Interessen berücksichtigt wurden. Im nächsten Schritt folgt eine Darstellung der Kontextbedingungen.

Kultusministerium als bildungspolitischer Akteur Das Kultusministerium des jeweiligen Bundeslandes ist für die Schulinspektion die oberste Dienstbehörde. Die hierarchische Stellung impliziert nach dem Sensemaking-Ansatz (→ Kapitel 5) jedoch nicht automatisch, dass Vorgaben des Ministeriums von Seiten der Schulinspektion so umgesetzt werden, wie es

vom Ministerium intendiert ist. So müssen formelle Hierarchieverhältnisse nicht unbedingt tatsächliche Machtverhältnisse widerspiegeln.

Generell zeigt sich, dass in allen Bundesländern ein inhaltlicher Austausch zwischen den Schulinspektionen sowie einzelnen Abteilungen der Kultusministerien stattfindet (z. B. Ing. 12_13; 10_34; 18_25). So werden das Ministerium bzw. Vertreter_innen aus dem Ministerium als ein Akteur genannt, der die Auseinandersetzung mit Inklusion durch die Schulinspektionen initiiert hat (Int. 1_23; 2_3; 8_25; 12_27; 7_21; 13_2; 17_15) bzw. durch eigenes Verhalten initiiierend wirkte (Int. 16_11). Des Weiteren tauschen sie sich über die konkreten Bewertungsinstrumente sowie über das tatsächliche Verfahren aus (Int. 9_11; 18_6; 7_17; 17_12; 16_14), z. B. über die Art und Weise, wie die Ergebnisse zurückgemeldet werden. Zusätzlich zeigt sich in den Daten, dass es in Bezug auf bestimmte Punkte konkrete Vorgaben oder Anweisungen von Seiten des Ministeriums gab. In anderen Fällen gab es zwar keine konkreten Vorgaben, aber die interviewten Personen antizipieren bei ihrem eigenen Handeln die Positionen des Ministeriums und bemühen sich darum, dass das eigene Handeln mit der Linie des Ministeriums konform ist.

Generell ist in Bezug auf das Kultusministerium zwischen der Entwicklung des gesamten Instrumentensets sowie der Entwicklung von Instrumenten zu Inklusion zu unterscheiden: So gibt es beispielsweise den Fall, dass generell ein inhaltlicher Austausch zu Bewertungsinstrumenten stattfindet, aber es wenige Hinweise zu Inklusion gab, weil das Kultusministerium selbst noch in der Findungsphase war (z. B. Int. 2_25, siehe auch 10_33). Es wird in einem anderen Fall aber auch die Vermutung geäußert, dass es politische Gründe sind, warum es keine konkreten Hinweise zu Inklusion gibt: *„Wenn die Voraussetzungen, also die rechtlichen Rahmenbedingungen, (.) äh, alle- alle Facetten offenlässt, dann erleben wir natürlich auch diese (.) Facetten. Also die Steuerung is wichtig (.) zu inklusiven- in diesem inklusiven Kontext. Und äh, da nehm ich (.) eher wahr, dass die Politik überlegt, wie se wiedergewählt werden kann, als dass jetzt inhaltlich gesteuert wird. Weil das sind andere- andere Faktorn. Im Augenblick is die große Angst vor der ähm-äh-a-a AfD zu spürn. Und die sind so vorsichtig, also mit Allem was sie tun und ja. [...,ap]. Dann die sprechen dann immer- die vom Kultusministerium von Leitplanken (.), ja? Aber die Leitplanken sind (.) schon ziemlich weit gefasst. Man könnte, und man sollte und.“* (Int. 14_50). Des Weiteren ergänzt sie später *„Die ham dann gesagt, äh-ähb-ähhh mit dem- mit dem Begriff Inklusion kann man keine Wahl gewinnen. Also die trennen dis sehr stark voneinander“* (Int. 14_50). Demnach hält das Kultusministerium in manchen Fällen konkrete Hinweise zurück, wie Inklusion operationalisiert werden könnte. Den Schulinspektionen wird somit Handlungsspielraum gelassen. Nichtsdestotrotz haben die Instrumenteentwickler_innen vor Augen, welche Erwartungen von Seiten der Kultusministerien formuliert werden: *„Das andere Problem, was wir, glaub ich, ham, is dass wir mit unseren Berichten natürlich nich nur Jubelbotschaften produziern und insofern auch politisch nich unbedingt vermarktbar sind. Das heißt also wir liefern im Grunde genommen (.) Erkenntnisse, die nich unbedingt gewollt sind. Und das heißt, (.) ähm, wir setzen uns an der Stelle also mehr oder weniger selber den Auftrag“* (Int. 13_13).

Des Weiteren gibt es auch den Fall, dass das Ministerium eigentlich keine inhaltlichen Vorgaben bei der Entwicklung von Instrumenten macht, aber in Bezug auf Inklusion aktiv wurde (z. B. Int. 17; 9_16; 12_3): *„Was die inhaltliche Seite anbelangt, also warum ist ein Kriterium in einem Tableau und dass die Schulen danach bemessen werden, da gab es eigentlich keine politischen Vorgaben. Erst nach fünf-sechs Jahren mit der Inklusion kam eine Vorgabe“* (Int. 17). Beispielsweise orientieren sich die Instrumenteentwickler_innen in ihrem Handeln daran, wie Inklusion offiziell auf ministerialer Ebene definiert wird (z. B. Int. 3_17; 9_36; 14_33; 7_21; 17_39): *„einfach mal*

schaue, wie wird Inklusion hier im Land kommuniziert. Was ist das? dann muss ich sagen. leider sind wir inzwischen an dem Punkt in [Bundesland], wo Inklusion gleich gesetzt wird mit gemeinsamen Unterricht und Betreuung von Kindern mit sonderpädagogischem Förderbedarf. Es wird zwar immer davon geredet, dass es den Index für Inklusion gibt, dass wir die Menschenrechtskonvention unterzeichnet haben, dass wir Schritte unternehmen müssen, um moderne Schulen zu entwickeln und und und. Aber der Begriff Inklusion wird explizit benutzt, um die GU-Kinder oder die I-Kinder oder wie auch immer. Also den Umgang mit diesen Kindern zu beschreiben. Ja. Und zu sagen: wie müsste der Unterricht aussehen, um sozusagen erfolgreich auf diese GU-Kinder beschulen zu können“ (Int. 9_37). Demnach wird in diesem Bundesland – wie in den meisten Fällen – von Seiten des Ministeriums ein enger Inklusionsbegriff vertreten. Dagegen zeichnet sich – wie bereits in Kapitel 7.3 dargestellt - auf Individualebene bei den interviewten Instrumenteentwickler_innen ein heterogenes Bild ab.

In diesem Zusammenhang zeigen sich Anknüpfungspunkte zu dem Phänomen des „street level bureaucrat“ von Lipsky (2010): Dieses basiert auf dem Ansatz, dass Akteure, die eine Vorgabe in der Praxis umsetzen sollen, vor der Herausforderung stehen, abstrakte Vorgaben zu konkretisieren und in ihre bisherigen Routinen zu integrieren. Hierbei verfügen sie über einen gewissen Handlungsspielraum. Dies ermöglicht es ihnen beispielsweise auch eigene Überzeugungen einzubringen. So beschreiben Wong und Anagnostopoulos (1998), wie Lehrkräfte mit rechtlichen Vorgaben umgehen, die eine stärkere Orientierung an Schulleistungstests implizieren. Dabei können sie nachweisen, dass die Lehrkräfte verschiedene Coping-Strategien entwickeln, um mit den Vorgaben der Bildungsadministration umzugehen. Beispielsweise implementieren viele Lehrkräfte die Vorgaben nur oberflächlich oder berücksichtigen die Vorgaben nur in einem Teil ihrer Unterrichtsstunden. Demnach wirken Vorgaben „von oben“ nur begrenzt.

Im Fall der Instrumenteentwickler_innen zeigt sich in Bezug auf den engen und weiten Inklusionsbegriff ein ähnliches Muster: So werden vereinzelt Coping-Strategien gewählt, um ein weites Inklusionsverständnis in den Instrumenten zu verankern: So argumentiert eine Interviewperson, dass ein weites Verständnis zwar nicht explizit mit dem Inklusionsbegriff in den Instrumenten verknüpft wird, aber bei dem Aspekt „Umgang mit Heterogenität“ mitgedacht wurde (z. B. Int. 3; 9): *„schon Elemente auch des Index oder eines breiteren Denkens von Inklusion. Aber wie gesagt nicht explizit ausgewiesen. Ja. Aufgrund dieses sehr engen Verständnisses von Inklusion vom Land“ (Int. 9). Demnach haben diese Instrumenteentwickler_innen den Eindruck, dass ihr eigenes Verständnis von Inklusion in den Instrumenten repräsentiert ist, nur nicht explizit mit dem Inklusionsbegriff verknüpft wird. Des Weiteren berichtet eine Interviewperson, dass sich die Schulinspektion als Ganzes zu einem weiten Inklusionsverständnis positioniert hätte: „ich-ich frag mich grad, ob das jetzige Kul-Kultusministerium irgenwo mal ne- (.) ähm-äh Definition hinterlegt hatte. Dis gibt einen sogenannten Aktionsplan zur Umsetzung. Bezieht sich aber auch immer nur auf die Umsetzung der Behindertenrechtskonvention. Also, die-die Definition is so, ja? Hier für Bundesland X [anonymisiert, ap]. Und wir ham son bisschen weiteren geschafft. Bisschen Freiheit genom.“ (Int. 15, siehe auch 9_44). Demnach zeigt sich hier, dass der vom Kultusministerium gesetzte Inklusionsbegriff ignoriert wird bzw. die eigenen Vorstellungen von Inklusion in die Instrumente Eingang finden. Um Konflikte zu vermeiden, werden die Instrumente in diesem Fall nicht als inklusionsorientiert gekennzeichnet.*

Schulische Vertreter Unter dem Begriff „schulische Vertreter“ werden im Folgenden unterschiedliche Personengruppen gefasst, die die Interessen der Praxis einbringen: Schultartvertreter, Vertretungen von Einzelschulen sowie Lobbyverbände, z. B. von Lehrer_innen wie der Philolo-

genverband (z. B. Int. 10_5). Viele dieser Akteure waren bei der Entwicklung der Instrumente repräsentiert. Des Weiteren wurden die Positionen und Sorgen der schulischen Vertreter bei der Entwicklung der Instrumente von Seiten der Inspektion, aber auch durch das Kultusministerium antizipiert, auch wenn diese nicht durch konkrete Personen vertreten waren.

Das Verhältnis zwischen Schulinspektionen und schulischen Vertretern ist maßgeblich davon geprägt, welche Funktionen die Inspektionen innehaben (nähere Ausführung zu den Funktionen von Schulinspektionen → Kapitel 8.1.3). Darüber hinaus ist zu berücksichtigen, dass die Schulinspektionen in der formellen Hierarchie über den Schulen bzw. den schulischen Vertretern anzusiedeln sind. Demnach wäre zu erwarten, dass die schulischen Vertreter bei der Entwicklung von Messinstrumenten keine allzu große Rolle spielen, weil es darum geht diese zu kontrollieren. Zur Illustration: Wenn die Kompetenzen von Schüler_innen im Rahmen von Klausuren abgefragt werden, sind diese bei der Entwicklung der Klausurfragen meistens nicht involviert. Diese Annahme zeigt sich jedoch nur begrenzt in den Daten: *„Alle Schulen müssen vorbereitet sein und die Frage, ob hier ein behindertes Kind ist oder nicht, spielt hier auch überhaupt keine Rolle. Und hier wären auch die Förder- und Realschulen [anonymisiert] wären da auch mitgegangen. Also die Hälfte der Schulen, wenn man soll. Aber die andere Hälfte nicht. Die einfach Angst gehabt haben - es würde jetzt ein Kriterium in dem Profil der Schule erscheinen, wo sie alle schlecht abschneiden, weil sie das alle gar nicht haben“* (Int. 17_18). Andere Interviewpartner wollen die Schulen zwar fordern – insbesondere, wenn diese noch nicht aktiv geworden sind (3_22), aber nicht überfordern (7_28; 12_31; 3_22; 8_16): *„Also zum einen, dass ähm dass die Schulen im Moment damit überfordert sind und dass es von daher unangemessen ist das überhaupt in den Blick zu nehmen. Also dass man sozusagen das trennen soll von der Schulinspektion [anonymisiert], dass es von daher sozusagen moralisch ist unsern Prüfauftrag ähm da in dieser Richtung zu sehen, dass ist die eine Begründung“* (7_28). Die Ängste werden nicht nur registriert, sondern auch bei der Entwicklung der Instrumente berücksichtigt. So gab es nur wenige Bundesländer, die den schulischen Vertretern bei der Beschreibung des Entwicklungsprozesses keine große Rolle beigemessen haben und es zeigte sich darüber hinaus, dass einzelne Inspektionen dafür bestraft wurden, wenn sie die Positionen der schulischen Vertreter zu wenig in ihrem Handeln berücksichtigt haben: *„Gut und (.). Ich denke mir, wir warn auch einige Jahre lang ähm, (.) ham wir aufn ziemlich hohen Ross gesessen, so nach dem Motto: Wir sind diejenigen, die wissen, was gute Schule ist. <lachen> Und jetzt dreht es sich eben und dann sagen die, die da bescheinigt bekommen haben: <Ihr macht keine gute Schule.> Die haben jetzt ma grade wieder mehr Macht und sagen: <So, jetzt zeigen wa den mal, was ne Harke is und legen die [die Schulinspektionen, ap] erstmal still.> (lachen) Das is einfach so.“¹* So zeigt sich in einer Vielzahl von Interviews, dass die (antizipierten) Positionen und Interessen von schulischen Vertretern auf unterschiedliche Art und Weise die Entwicklung von Instrumenten entscheidend mitprägen. Dies spiegelt sich beispielsweise darin wider, dass die Notwendigkeit formuliert wird, dass das Verfahren der Inspektion und die Instrumente bei den Schulen Akzeptanz erfahren müssten (Int. 2_9; 8_36; 18_15; 5_38; 11_17). So hatte beispielsweise ein Bundesland zu Beginn nach eigener Einschätzung ein „sanftes Inspektionsverfahren“: *„Dass man sehr vorsichtig auf die Schulen zugehen wollte und sehr offen und sehr wertschätzend Ich denke, dass sich dieser eher sanfte Einstieg, den wir hier in [Name des Bundeslandes, ap] gewählt haben, durchaus bewährt hat. Also wir hatten ja jetzt keine großen Auf und Abs - also das Thema [die Inspektion von Schulen, ap] ist jetzt quasi stabil seit [anonymisiert, ap] Jahren“* (Int. 2_9). Demnach werden die Positionen von Seiten der

1 Aus Datenschutzgründen keine genauere Angaben zum Interviewpartner.

schulischen Vertreter berücksichtigt, um zu erreichen, dass das Inspektionsverfahren als solches von ihnen Akzeptanz erfährt (nähere Ausführung siehe Piezunka 2019).

Des Weiteren wird die Nicht-Einführung der Schulrankings damit begründet: *„und das war auch von Anfang an klar. Es gibt kein Ranking. Das war eine Zusage von Ministeriumsseite her, weil die - die ganzen Verbände interveniert haben, dass sie so etwas nicht haben wollen. das ist das eine und auch die Schulen haben gesagt - sie wollen so etwas eigentlich garnicht gern. Alle haben die Befürchtung gehabt - das führt dann zu einem Schulranking und drum haben wir das von Anfang an garnicht vorgehabt“* (Int. 17_12). Dennoch erhalten die Mitarbeiter_innen von Schulinspektionen von einzelnen Schulen die Rückmeldung, dass diese sich gerne im Vergleich zu anderen Schulen verorten würden (Int. 12_12) bzw. in der Vergangenheit versucht haben, dies mit den vorhandenen Daten zu tun (Int. 5_24). Generell werden die schulischen Vertreter jedoch als ein Akteur wahrgenommen, der sich gegen Schulrankings positioniert.

Am Beispiel der Rankings wird die Machtposition der schulischen Vertreter sichtbar, wenn der Fall der deutschen Schulinspektionen mit anderen Fällen verglichen wird, z. B. mit den Hochschulrankings im US-amerikanischen Bildungssystem. So weisen Espeland und Sauer (2016) darauf hin, dass die Hochschulrankings in den USA von dem Magazin „US News & World Report“ eingeführt wurden. Einige Hochschulen versuchten sich dem Ranking zu entziehen. Dies führte dazu, dass die Daten dieser Hochschulen von dem Magazin einfach geschätzt wurden und hierfür auf Schätzwerte zurückgegriffen wurde, die von Seiten der betroffenen Hochschulen als schädigend empfunden wurden (vgl. Sauder und Espeland 2009, S. 76). Dies hatte zur Folge, dass diese Hochschulen an weiteren Erhebungen aktiv teilnahmen, um Einfluss auf die Daten nehmen zu können. Demnach stellte sich hier gar nicht die Frage, inwiefern es ein Ranking geben würde, und die Verhandlungsposition der Hochschulen war schlecht, da sie sich aufgrund der breiten öffentlichen Rezeption der Rankings diesen nicht entziehen konnten. In diesem Zusammenhang stellt sich die Frage, wie es sich erklären lässt, dass sich Rankings bei der deutschen Bildungsadministration nicht durchsetzen konnten. Wie bereits oben erwähnt, könnte ein Aspekt sein, dass die schulischen Vertreter im deutschen Bildungssystem eine mächtigere Position innehaben. Des Weiteren zeigt sich, dass Rankings im deutschen Bildungssystem – im Vergleich zum US-amerikanischen System, welches im Buch von Espeland und Sauder beschrieben wird – im Allgemeinen eine geringere Rolle spielen. Mögliche Ursachen können nach einer These von Espeland sein, dass es im Vergleich zwischen USA und Deutschland weniger Wettbewerb zwischen Schulen gibt und dass der Besuch von bestimmten Schulen bzw. bestimmten Hochschulen in Deutschland für den weiteren Bildungsverlauf weniger entscheidend ist als in den USA. Da die Schul- bzw. Hochschulwahl in den USA für den weiteren Lebensverlauf von hoher Relevanz ist, ist auch die Auswahl der „richtigen“ Hochschule von größerer Bedeutung. Um diese Entscheidung treffen zu können, greifen Eltern und Schüler_innen in den USA aufgrund von Unsicherheit auf Rankings zurück. Demnach gibt es in Deutschland von Seiten der Eltern und Schüler_innen einen geringeren Bedarf an Rankings.

Die Relevanz von schulischen Vertreter zeigt sich auch im Hinblick auf die inklusionsorientierten Instrumente, z. B. als es darum ging zu entscheiden, welche Anforderungen für machbar gehalten werden (Int. 1_41; 15_23) und inwiefern die Daten generell für die Schulen von Relevanz sind (Int. 9_17; 16_9). Darüber hinaus orientieren sich die Schulinspektionen an schulischen Vertretern, da es ihnen ein Anliegen war, durch ihre Bewertungsinstrumente Themen abzudecken, die für die Schulen aktuell von Relevanz sind (Int. 2_10; 10_18). Demnach sind die Schulinspektionen zwar in der Hierarchie formell über den Schulen gestellt. Dennoch

zeigt sich in den Daten, dass die Positionen der schulischen Vertreter bei der Entwicklung von Bewertungsinstrumenten eine maßgebliche Rolle gespielt haben. So werden von einzelnen Instrumenteentwickler_innen Zusammenhänge zwischen der Akzeptanz durch schulische Vertreter und dem Existenzstatus von Schulinspektionen gesehen.

Schulaufsicht Im Kapitel zu Schulinspektionen (→ Kapitel 3) wurde das Verhältnis zwischen Schulinspektionen und Schulaufsicht näher beschrieben: So ist eine Zusammenarbeit notwendig, da in fast allen Bundesländern die Schulaufsicht dafür verantwortlich ist, zu überprüfen, inwiefern die Schulen nach dem Besuch der Schulinspektion aktiv werden (z. B. Int. 10_9), z. B. indem Zielvereinbarungen zwischen Schulaufsicht und Schulinspektion getroffen werden. In einem Interview gibt es jedoch die Andeutung, dass dies nicht immer funktioniert: *„Die ursprüngliche Idee war ja ähm wir machen die Rechenschaftslegung und dann arbeitet die Schulaufsicht mit ihrem Beratungs- Unterstützungssystem weiter - dabei die Schule zu beraten. Ähm das funktioniert nicht so gut, wie man sich das vorstellt. Also in der Theorie geht das, in der Praxis funktioniert einfach so nicht, weil die Dezernenten aus welchen Gründen auch immer, Arbeitsbelastung, zu viele Schulen, nicht in ausreichender Weise dazu kommen, sich mit den Schulen nach der Inspektion so auseinanderzusetzen, dass sie die da abholen, wo sie gerade sind und dann entsprechend weiter begleiten“* (Int. 10_9).

Hierbei zeigt sich ein Phänomen, welches Heinrich (2015) bereits herausgearbeitet hat: Die Schulinspektion wird von Seiten der Schulaufsicht als ein Akteur wahrgenommen, der den eigenen Kompetenzbereich bedrängt. Daher ist es den Interviewpersonen ein Anliegen, den eigenen Kompetenzbereich deutlich von dem der Schulaufsicht abzugrenzen: *„Auch in der Findung der Rolle, die ja ganz neu war, und gerade so in der Platzierung zwischen Schulaufsicht und Landesinstitut für Beratung, mussten wir unsere Nische suchen, ohne den anderen Beteiligten zu vermitteln, wir kommen in euren Bereich“* (Int. 16_19, siehe auch 12_18). In diesem Zusammenhang spielt auch eine Rolle, was als primäre Funktion von Schulinspektionen angesehen wird. So sagt eine Interviewperson bei der Beschreibung der Schulinspektion als Impulsgeber für Schulen: *„Ich sage jetzt mal böse: wir sind ja eigentlich die Guten, die versuchen mit den Schulen inhaltlich zu arbeiten und ähm ja im Sinne von Kommunikation die Handlungsfelder zu entwickeln“* (Int. 9). Dagegen wird die Rolle der Schulaufsicht von dieser Interviewperson stärker in der Kontrollfunktion gesehen (Int. 9).

Im Hinblick auf die Instrumente spielt die Schulaufsicht eine Rolle, da sie bei der Entwicklung miteinbezogen wurde (z. B. Int. 10_6; 16_17; 18_5; 17_57), z. B. als Mitglieder in den Arbeitsgruppen. Die Anerkennung des schulaufsichtlichen Kompetenzbereichs zeigt sich in vielen Fällen auch darin, dass keine Aspekte abgefragt werden, mit denen nicht nur die Schule, sondern auch die Schulaufsicht evaluiert werden würden, z. B. Rahmenbedingungen an der Schule (Int. 17_9; 10_37) oder auch die Frage, woran die Schule seit der letzten Schulinspektion gearbeitet hat: *„Wir wollten nämlich bei der Revision gerne ein Kriterium aufnehmen äh und zwar in der Gestalt und zu prüfen, was ist aus der Schulinspektion [anonymisiert, ap] geworden, was hat die Schule gemacht und das wurde abgelehnt und zwar deshalb, weil es eine Aktion ist, die in Kooperation mit der Schulaufsicht ist und es uns nicht zusteht auch die Schulaufsicht zu evaluieren. Also da war der Punkt. Das ist eigentlich eine Bewertung der Schulaufsicht - das machen wir nicht. Letztlich geht es aber garnicht, weil die Schulen natürlich bei dem Kriterium [anonymisiert, ap] Qualitätsentwicklung Dinge einbringen, die sie zusammen mit der Schulaufsicht machen. Also letztendlich kann man das garnicht verhindern“* (17_58; siehe auch 1_39). Aus Rücksichtnahme gegenüber der Schulaufsicht ist es in diesem Fall für die Schulinspektion

schwierig zu überprüfen, inwiefern ihre letzte Inspektion Wirkung gezeigt hat. Ein weiterer Aspekt wird von Botz et al. angesprochen, nämlich die Frage, „wer evaluiert, ob die vom Land gebotenen Mittel ausreichend sind“ (Bolz et al. 2011, S. 31). Indem die Schulinspektionen sich darum bemühen, nicht in den Kompetenzbereich anderer Akteure einzudringen, gibt es keine Kontrolle darüber, inwiefern die Rahmenbedingungen vorhanden sind, damit Schulen in der Lage sind, bestimmte Anforderungen zu erfüllen. Bezüglich Inklusion bedeutet dies, dass durch die Daten der Schulinspektion beispielsweise nicht ermittelt werden kann, ob die Schulen die Ressourcen erhalten, die ihnen für die Umsetzung von Inklusion zustehen. Abschließend lässt sich im Hinblick auf die Schulaufsicht festhalten, dass in den Interviews kaum Aussagen darüber getroffen werden, welche inhaltlichen Positionierungen zu den einzelnen Instrumenten von ihnen ausgehen und wie diese konkret mitwirken.

Im Austausch mit anderen Schulinspektionen Bei der Entwicklung von Bewertungsinstrumenten berücksichtigen die Instrumenteentwickler_innen auch, was andere Schulinspektionen machen. So hospitieren die Schulinspektionen gegenseitig (z. B. Int. 1_5; 2_6; 8_12; 11_10). Für die Inspektionen ist nicht nur von Interesse, was in anderen Bundesländern passiert (Int. 1_5; 8_3; 12_22; 18_32; 5_15; 14_43), sondern auch was im Ausland passiert: So findet beispielsweise ein Austausch mit England, Schottland, Niederlande, Österreich, Nordirland, Litauen und Portugal (z. B. Int. 8_2; 14_43; 5_15; 10_8) statt. Beim inhaltlichen Austausch geht es unter anderem darum zu erfahren, welche Kriterien und Messinstrumente (z. B. Int. 1_6) genutzt werden und inwiefern es Unterschiede in Bezug auf das Verfahren oder primäre Funktionen (Int.10_8) gibt. So wurde einer Schulinspektion beispielsweise von anderen Bundesländern abgeraten, Inklusion als „Add-On“ in die Bewertungsinstrumente zu integrieren (Int. 2_20).

Zwischen den Bundesländern unterscheidet sich, wie intensiv der Kontakt ist: Beispielsweise haben sich bestimmte Bundesländer zu einem Netzwerk zusammengeschlossen, um gegenseitig stattfindende Hospitationen zu institutionalisieren. Jedoch gibt es auch auf informeller Ebene Austausch z. B. über bestehende Kontakte oder geografische Nähe. Generell entsteht der Eindruck, dass fast alle Schulinspektionen im direkten Austausch mit anderen Bundesländern sind „Also der Wille [zum Austausch, ap] ist schon da, ne äh (...) nur Bundesland X meldet sich nicht“ (Int. 14_43).

Der Austausch wird insofern als positiv bewertet, als dass alle von den Erfahrungen der anderen profitieren können: „Jetzt habe ich einfach daran gedacht, dass wir in den nächsten Monaten eher informelle, eher auf einer kollegialen Ebene, auf andere Inspektorate zu gehen werden, um einmal das Gespräch zu suchen [...].ap]. Wird das interessant mal zu hören, gibt es vielleicht neue Schwerpunkte die gesetzt werden, ähm gibt es neue interessante Verfahrensweisen, Instrumente, können wir sozusagen Honig saugen aus Erfahrungen und Überlegungen der Kooperationspartner.“ (Int. 8_33). Der Austausch bietet die Möglichkeit von Best-Practice-Instrumenten und Verfahren zu profitieren, jedoch weist ein anderer Interviewpartner daraufhin, dass es nur eine begrenzte Übertragbarkeit gibt, weil die „Möglichkeiten und Gegebenheiten“ (Int. 12_22) zwischen den Bundesländern voneinander abweichen.

Forscher_innen oder Forschungsergebnisse Im Folgenden geht es nun darum, inwiefern wissenschaftliche Forschung bei der Entwicklung von Instrumenten im Allgemeinen sowie in Bezug auf Inklusion berücksichtigt wird.

- Nutzung von Forschung für die Entwicklung von Instrumenten im Allgemeinen: Einige Schulinspektionen stellten durch ihre Personalauswahl sicher, dass wissenschaftliche Expertise im eigenen Team vorhanden ist (z. B. Int. 13_19; 4). Eine weitere Form der Auseinandersetzung mit wissenschaftlicher Forschung findet statt, indem Mitarbeiter_innen der Schulinspektionen sich bei der Entwicklung von Instrumenten intensiv mit der vorhandenen wissenschaftlichen Literatur auseinandersetzen (z. B. Int. 1_11; 2_7; 4_4; 10_25). So antwortet ein Instrumententwickler auf die Frage, wie sie zu den Bewertungsinstrumenten gekommen sind: *„Wir haben uns da durch die Literatur durchgefressen und haben alles, was es dazu gibt. äh. aufgearbeitet und versucht aus dem raus Kriterien, Anforderungen - also - ich weiß nicht, ob die anderen es auch so haben“* (Int. 17_41).

Hierbei fällt auf, dass die Namen von drei Wissenschaftlern immer wieder genannt werden: Andreas Helmke (Int. 12_2; 8_4; 6_13; 17_7), Hilbert Meyer (Int. 12_2; 8_4; 9_9; 6_13; 11_8) sowie John Hattie (Int. 3_5; 8_4; 10_2; 11_8): *„und hier haben eigentlich glaube ich die meisten Länder und auch wir ursprünglich zurückgegriffen auf das was Helmke und Meyer und andere entwickelt haben und ähm Sie können sich denken, aus meiner Sicht jedenfalls, das natürlich jetzt in den letzten zwei drei Jahren ganz stark die Forschung von Hattie Einfluss genommen haben und jetzt auch häufig diesen Begriff lernwirksamer Unterricht verwenden. Das ist natürlich dann die Grundlage gewesen“* (Int. 8_6) oder ein anderer Interviewpartner erklärt *„Wir haben uns da im zweiten Teil eigentlich sehr stark an Helmke gehalten, der war im Grunde genommen eigentlich der Mentor für uns“* (Int. 17).

Eine mögliche Erklärung für den häufigen Verweis auf diese drei Autoren ist, dass diese in der fachlich informierten Öffentlichkeit breit rezipiert werden. Beispielsweise hat die Wochenzeitung „Die Zeit“ sowie andere Zeitungen immer wieder über die Publikation „Lernen sichtbar machen“ von John Hattie berichtet oder die Publikation zu gutem Unterricht von Hilbert Meyer erschien im Jahre 2016 bereits in elfter Auflage. Die Beobachtung, dass sich Praktiker an breit rezipierten und leicht zugänglichen Publikationen orientieren, zeigt sich auch in anderen empirischen Studien (z. B. Penuel et al. 2018). So beschreiben Penuel et al. diese Form von Publikationen als „How-to-guides“, weil sie konkrete Handlungsanweisungen für die Praxis geben. So spielt die Praxisnähe bzw. –relevanz eine wichtige Rolle bei der Frage, ob wissenschaftliche Forschungsergebnisse in der Praxis genutzt werden.

Auch wenn sich viele der Interviewpartner an diesen drei Autoren orientieren, gibt es auch eine Interviewperson, die sich explizit von Hattie abgrenzt und bei der Frage nach relevanter Literatur auf die Schuleffektivitätsforschung verweist: *„Was ist Stand der empirischen Forschung? Also, weil im Grunde genommen is natürlich unser Anspruch, wir müssen hier irgendwie State Of The Art das - den aktuellen Stand der Schuleffektivitätsforschung abbilden. (.) Ähm (.), Hattie war nicht unsere Bibel <lachen>“* (Int. 13_20). Jene abweichende Positionierung lässt sich möglicherweise dadurch erklären, dass die befragte Person selbst über einen längeren Zeitraum im wissenschaftlichen Bereich tätig war.

Abschließend lässt sich festhalten, dass viele der interviewten Instrumententwickler_innen auf Helmke, Meyer und Hattie verweisen. Zusätzlich gibt es noch einzelne Verweise auf die Schuleffektivitätsforschung (Int. 16_5; 13_20) sowie auf wissenschaftliche Literatur aus der Demokratiepädagogik (Int. 9_9) und Ditton (Int. 17_3).

Neben der Rezeption von Literatur wurden Wissenschaftler_innen zum persönlichen Austausch eingeladen, um Rückmeldung auf die Instrumente zu geben (z. B. Int. 8_10; 6_12). Bei der Nachfrage, wie es zu der jeweiligen Auswahl der Wissenschaftler_innen kam, wurde auf drei Begründungen zurückgegriffen: Geographische Nähe, d. h. die jeweiligen Wissen-

schaffler_innen waren z. B. an einer Hochschule des jeweiligen Bundeslandes angestellt (Int. 2_7; 18_5; 6_13). Des Weiteren kontaktierten sie Wissenschaftler_innen, die sie für bekannt und für den Forschungsbereich für relevant hielten (Int. 6_13; 10_26): „Also wenn ich jemandem zum Unterricht frage, dann fallen mir drei, vier Namen ein und viel mehr gibt's dann auch nicht. Und dann fragt man bei dreien oder vieren an und zwei können dann zufällig. Also es hängt auch tatsächlich mit Zufälligkeit zusammen“ (Int. 10_26). Darüber hinaus griffen Schulinspektionen auch auf bereits bestehende Kontakte zurück, die zwischen den Wissenschaftler_innen und dem jeweiligen Kultusministerium (Int. 6_14; 10_26) bzw. einzelnen Personen (Int. 8_4) bestanden.

- Nutzung von Forschung in Bezug auf Inklusion: Die Frage, welche Literatur für die Entwicklung von Instrumenten zu Inklusion genutzt wurde, wird sehr häufig auf den „Index für Inklusion“ verwiesen (Int. 14_9; 15_25; 3_5; 9_42; 14_9; 16_28; 13_4): „Es gibt Bewertungsinstrumente [anonymisiert, ap], das habe ich entwickelt, das liegt in der Schublade. [...,ap] Also wie gesagt, es liegt noch in der Schublade. Das hebt auf den Index für Inklusion ab. Dass ein Stück wird hinterfragt oder hinterfragen würde, wenn wir das zum Einsatz bringen. Welches Weltbild habt ihr, welches Pädagogikbild habt ihr im Kopf, welche Strukturen habt ihr entwickelt, um ne Willkommenskultur anbieten zu können. Wie wirkt sich das im Unterricht aus? Also diese drei groben Blöcke des Index“ (Int. 9). Neben dem Index für Inklusion wird das Aargauer Bewertungsraster (Int. 14_9; 15_25) genannt sowie die Arbeiten von Feuser (Int. 9_2). Auch wenn viele auf den Index für Inklusion verweisen, hat sich dieser in der konkreten Anwendung für die Instrumenteentwickler_innen nur in Einzelfällen (z. B. Int. 9) als nützlich erwiesen. Dies lässt sich dadurch erklären, dass dieser als Selbstevaluationsinstrument auf die reflexive Auseinandersetzung in der Einzelschule abzielt (Int. 14_12). Darüber hinaus steht der Index aufgrund fehlender Trennschärfe und Nicht-Messbarkeit durch empirische Methoden in der Kritik (Int. 13_5).

Des Weiteren gab es im Hinblick auf Inklusionsforschung in den Interviews Aussagen, wonach es an Literatur noch fehlt: „da ist Forschung ja - sicherlich gibt es da auch Literatur. Aber das ist ja alles sehr wenig entwickelt - sage ich mal“ (Int. 2_21; siehe auch 4_4). Darüber hinaus kritisiert eine Interviewperson die fehlende empirische Ausrichtung von Inklusionsforscher_innen: „Ja, wir haben uns- ähm, ja wir ham uns [Name eines Forschers] eingeladen, irgendwie. Das war nachdem ich irgendwie den erlebt hatte bei na Tagung und gemerkt hab: Oh Gott sei Dank, es gibt irgendwie Leute, die Inklusion betreiben und trotzdem empirisch ticken und nich ideologisch verbrämt sind, hab ich gesagt, so, den brauchen wir jetzt!“ (Int. 13_37).

Zusammenfassend zeigt sich, dass im tatsächlichen Handeln immer wieder die Nähe zur Forschung gesucht wird – sei es durch persönliche Kontakte oder durch Literaturrecherche. Hierbei zeigt sich als Phänomen, dass bei der Entscheidung darüber, wessen Forschungsarbeiten berücksichtigt werden, unter anderem die bestehende Kontakte und geographische Nähe eine Rolle spielen. In Bezug auf Inklusion wird jedoch deutlich, dass die bestehende Forschung zu Inklusion als nicht hilfreich erachtet wird, um daran anknüpfend Bewertungsinstrumente zu Inklusion zu entwickeln. Es stellt sich generell auch die Frage, welche Implikationen der Bezug zur Forschung für die konkrete Entwicklung von Instrumenten hat. So wird im weiteren Verlauf dieser Arbeit diskutiert, inwiefern der Verweis auf Forschung möglicherweise auch dazu genutzt wird, um das eigene Handeln retrospektiv zu legitimieren (nähere Ausführung → Abschnitt 9.2).

Schüler_innen und Eltern Auf die Frage, welche Akteure bei der Entwicklung der Instrumente eine Rolle gespielt haben, wurden nur selten Eltern oder Schüler_innen genannt. In einem Fall wurden sie zum Ende des Entwicklungsprozesses einbezogen (Int. 10_5). Als Wunsch von Seiten der Eltern wird genannt, dass die Inspektion stärker Aspekte wie Schulgröße, Klassengröße usw. berücksichtigen sollten. Diesem Wunsch möchte die befragte Interviewperson nicht entsprechen. Dies wird damit begründet, dass diese Aspekte nicht im Handlungsspielraum der Einzelschule liegen: *„Das war manchmal ein bisschen schwierig Leuten klar zu machen, dass das Strukturgegebenheiten sind für die die Schulen nicht verantwortlich gemacht werden können“* (17_9). Ausgehend von der Perspektive, dass die Qualität von Schule gemessen werden soll, um sicherzustellen, dass die Teilhabe von Schüler_innen gewährleistet wird, ist es überraschend, dass diese – im Gegensatz zu anderen schulischen Vertretern – kaum angehört werden.

Weitere Akteure Neben den oben aufgeführten Akteuren werden nur vereinzelt noch andere Akteure genannt. So berichtet ein Bundesland darüber, dass es in Bezug auf Inklusion einen inhaltlichen Austausch mit der Behindertenbeauftragten des jeweiligen Bundeslandes gegeben habe. Ein anderer Interviewpartner erzählt, dass bei der Entwicklung der Instrumente die Studienseminare miteinbezogen wurden (Int. 11_10). Demnach scheinen weitere Akteure, wie Interessensverbände von Menschen mit diagnostizierter Behinderung oder Menschen mit Migrationsgeschichte, keine relevante Rolle gespielt zu haben.

Die Analyse der Interviews zeigt, dass die Interviewpersonen die Interessen und Positionierungen von unterschiedlichen Akteuren berücksichtigen. So zeigt sich das Bedürfnis, dass die Instrumenteentwickler_innen den Erwartungen von verschiedenen Akteuren gerecht werden möchten: Auch wenn die Relevanz von einzelnen Akteuren zwischen den Bundesländern variiert, spielen die Kultusministerien, schulische Vertreter sowie Forscher_innen in fast allen Fällen eine entscheidende Rolle. Hierbei gibt es jedoch Unterschiede darin, wie diese sich positionieren bzw. wie deren Positionen von Seiten der Instrumenteentwickler_innen antizipiert werden. Beispielsweise gibt es von Seiten der Ministerien unterschiedliche Vorgaben bezüglich des Inklusionsbegriffs. Besonders hervorzuheben sind auch die schulischen Vertreter: Auch hier wird von Seiten der Schulinspektion Zusammenarbeit angestrebt, um die Wirkung des Verfahrens zu erhöhen: *„Sondern es ist eher so, ähm, dass wir versuchen, weil wir wissen wir müssen mit den anderen und wollen auch mit den anderen Akteuren zusammenarbeiten, wenn wir wirkungsvoll sein wollen, dass wir die daran beteiligen“* (Int. 16_15). Während die Ministerien und schulische Vertreter häufig genannt werden, scheinen Eltern, Schüler_innen und Lobbyverbände, die marginalisierte Gruppen repräsentieren, eine geringe Rolle zu spielen. Neben konkreten Akteursgruppen werden in den Interviews unterschiedliche Kontextbedingungen thematisiert, die bei der Entwicklung der Instrumente prägend wirkten, z. B. öffentliche Debatten, Umsetzung von Inklusion im jeweiligen Bundesland oder finanzielle und zeitliche Kapazitäten.

Öffentliche Debatten im jeweiligen Bundesland Wenn die Interviewpartner über die Entwicklung von Instrumenten zu Inklusion sprechen, verweisen einige auf die öffentlichen Debatten, die zu Inklusion stattgefunden haben. Dabei geht es darum, wie das Thema in der Öffentlichkeit diskutiert wird und welche Positionen den Diskurs dominieren. So hat die öffentliche Debatte – neben anderen Rahmenbedingungen – dazu geführt, die Auseinandersetzung mit Inklusion bei den Instrumenteentwickler_innen zu initiieren: *„Aber es war, es war halt plötzlich überall Thema in allen Zeitungen, die Politiker, die Bildungspolitiker sprachen dann überall davon ähm das war klar, dass [Bundesland] in der Pflicht ist, jetzt endlich mal die ähm,*

die Richtlinien umzusetzen“ (Int. 10_33). Des Weiteren ging es auch um die Frage, wie sich die Schulinspektion als Akteur innerhalb der Debatte positioniert (z. B. Int. 1_20). Bei einer Interviewperson ging es beispielsweise darum, ob die jeweilige Schulinspektion den engen Inklusionsbegriff übernimmt, der den Diskurs dominiert, und inwiefern sich die Schulinspektion gegen die kritische Haltung gegenüber Inklusion positioniert. Bei der näheren Betrachtung dessen, wie Inklusion in den jeweiligen Fällen gedeutet wurde, entsteht der Eindruck, dass die öffentlichen Debatten zwar wahrgenommen und verfolgt wurde, aber inhaltlich nicht handlungsleitend wirkten.

Umsetzung von Inklusion im jeweiligen Bundesland In Bezug auf die Umsetzung von Inklusion an den Schulen des jeweiligen Bundeslands wägen die Instrumenteentwickler_innen ab: Was ist machbar bzw. welchen Anforderungen können wir gegenüber Schulen formulieren (z. B. Int. 12_31; 3_7): *„Ja, wir ham (...) Neugier darauf, müssen natürlich auch immer anerkennen, wenn-wenn-wenn-wenn in Schulen Lehrkräfte diese Belastungen, diesen Zeitdruck äh anführen, warum sie bestimmte Dinge nur schwer umsetzen können. Dass wir darauf natürlich irgendwo auch Rücksicht nehmen und die Situation der Lehrkräfte äh anerkennen. Gleichwohl äh versuchen natürlich dann immer darauf hinzuweisen, ja das sollte doch aber so getan werden, das ist doch auch die Zielsetzung der Schule und jede Schule, die sich ähm als gute Schule versteht, versucht ja zu fördern. Und hier müssen son bisschen die Waage halten zwischen dem, was wir als äh Anspruch an gute Schule verstehen und dem, was ähm-äh auch von den Kollegen zu verlangen ist“* (Int. 12_31).

So versuchen die Interviewpersonen anhand ihrer Beobachtungen als Inspektor_innen und Erfahrungen als Lehrkräfte sowie im Gespräch mit schulischen Vertretern die Situation im eigenen Bundesland einzuschätzen: *„Ich denke die darf man nicht komplett losgelöst betrachten von dem Bedingungsgefüge in dem sich einerseits die Länder befinden, von dem Rechtsgefüge in dem sich die Länder befinden - das kann man als orientierende Schnitte betrachten und sagen: Die haben die Werte, die haben die Werte und letztendlich muss sich dann doch immer jedes Land, jedes Bundesland seine eigenen Marken setzen - im Sinne von - es mag zwar internationale Durchschnittswerte geben, aber wo setze ich meinen eigenen Fokus hin, welche Problematiken habe ich gerade zu bewältigen“* (Int. 11_15).

Bei der Einschätzung zur Umsetzung von Inklusion zeigt sich bundesländerübergreifend ein ähnliches Bild: So befinden sich Schulen zwar auf dem Weg (Int. 8_16), aber die Umsetzung von Inklusion erfordert von Seiten der Schule viele Veränderungen (Int. 11_28; 12_31) und viele Schulen seien noch nicht so weit (Int. 3_8; 12_36; 17_17): *„Ja, dass wir einfach ganz deutlich gemerkt haben, was können wir von Schulen erwarten, die zum Beispiel noch garnicht so weit sind. Ja, also. Ich glaube, weil das schon bei uns war - gemessen auch an anderen Bundesländern – [unser Bundesland, ap] so ein bißchen hinterher hinkte - so fühlte sich das immer an. Und auf einmal ging das alles Schlag auf Schlag, ja? und von der Theorie - und trotzdem kamen die Schulen nicht so ganz nach. Ja, also? [...,ap] und da war die Frage, wo sind unsere Erwartungen letztendlich da“* (Int. 3_8).

Dabei werden als Ursachen unter anderem die fehlende Haltung der relevanten Akteure (Int. 9_48; 14_36) sowie die fehlenden Rahmenbedingungen, z. B. Ressourcenausstattung oder exkludierende Strukturen im Bildungssystem (Int. 14_36, 15_29; 5_31), genannt. Demnach prägt die Umsetzung von Inklusion im jeweiligen Bundesland die Auseinandersetzung mit der Frage, was man bezüglich Inklusion von Seiten der Schule erwarten darf (Int. 1_38; 2_27;

16_39) und wirkt sich dementsprechend auch darauf aus, welche Anforderungen gegenüber Schulen formuliert werden.

Die hier beschriebene Kontextbedingung „Umsetzung von Inklusion im eigenen Bundesland“ weist Überschneidungen zu den Interessen der schulischen Vertreter (→ Abschnitt 8.1.3) sowie der eigenen Praxiserfahrung von einzelnen Instrumenteentwickler_innen auf (→ Kapitel 8.1.2): Wenn die Instrumenteentwickler_innen eine Entscheidung darüber treffen, wie viel sie von den Schulen erwarten können, beziehen sie die Rückmeldungen der schulischen Vertreter ein, aber auch – wenn vorhanden –, welche Erfahrungen die Instrumenteentwickler_innen selbst als Lehrkräfte im gemeinsamen Unterricht bzw. bei ihren Beobachtungen als Inspektor_innen gemacht haben.

Schulqualitätsrahmen Bei der Entwicklung von Messinstrumenten haben auch die Schulqualitätsrahmen den Prozess mitgeprägt. Hierbei ist jedoch zu bedenken, dass die Schulqualitätsrahmen in den einzelnen Bundesländern zu unterschiedlichen Zeitpunkten entstanden sind: Beispielsweise gibt es mehrere Fälle, wo die Schulqualitätsrahmen bereits existierten, als die Bewertungsinstrumente entwickelt wurden. Es gibt auch den Fall, dass die Bewertungsinstrumente für den ersten Zyklus der Schulinspektion bereits vorhanden waren, dann der Schulqualitätsrahmen entstanden ist und dieser wiederum als Grundlage für diejenigen Messinstrumente diente, die für den zweiten Zyklus entwickelt wurden. Demnach prägte der Schulqualitätsrahmen die Entwicklung der Instrumente, aber in manchen Bundesländern viceversa. Darüber hinaus zeigt sich in den Interviews, dass es auch personelle Überschneidungen darin gibt, wer für die inhaltliche Gestaltung des Schulqualitätsrahmens bzw. für die Bewertungsinstrumente verantwortlich ist.

In den Interviews verweisen mehrere Personen auf den Schulqualitätsrahmen, um ihr eigenes Handeln zu legitimieren (nähere Ausführung → Abschnitt 9.2). Außerdem dienen die Schulqualitätsrahmen in vielen Bundesländern als Grundlage für die Bewertungsinstrumente (Int. 2_2; 14_26; 18_3; 11_3; 5_8; 7.2; 13_18), z. B. „das ist für uns das Maß aller Dinge“ (Int. 15_12). Die Interviewpersonen unterscheiden sich jedoch darin, welche Relevanz sie dem Schulqualitätsrahmen zuschreiben. So werden verschiedene Begrifflichkeiten verwendet, z. B. Ableitung (Int. 2); Messlatte (Int. 15_12); als Orientierung (Int. 18_3), „als Basis nicht nur als Orientierung“ (Int. 5_8). So variiert in den Interviews die Einschätzung darüber, wie nah die Instrumente am Schulqualitätsrahmen geblieben sind. Viele Interviewpersonen geben an, dass nicht alle Aspekte aus dem Orientierungsrahmen bei den Instrumenten übernommen werden können (Int. 2_2; 14_26; 5_3; 6_4; 13_23; 16_5). Als Begründung wird hierfür beispielsweise genannt, dass zum einen nicht alle Aspekte aufgrund von finanziellen und zeitlichen Beschränkungen evaluiert werden können (z. B. Int. 16_5) und zum anderen, dass manche Aspekte nicht durch die Schulinspektionen messbar bzw. beobachtbar seien (Int. 6_4; 16_5; 5_3). Des Weiteren erfordert die Übersetzung in empirisch überprüfbare Indikatoren, dass Umformulierungen stattfinden, z. B. Konkretisierungen (Int. 5_3; 16_8). Die Auswahl an inhaltlichen Aspekten, die Übersetzung in mess- und beobachtbare Indikatoren sowie die Konkretisierungen führen dazu, dass der Schulqualitätsrahmen für den Handlungsbereich der Schulinspektion rekontextualisiert wird. In diesem Prozess finden darüber hinaus nochmal diverse Feedbackschleifen innerhalb des Hauses statt, sowie Anpassungen, die sich durch Abgleichungen mit Forschung und Inspektionspraxis ergeben (Int. 2_2; 9_5). Demnach lässt sich festhalten, dass dem Schulqualitätsrahmen nach Aussagen der Interviewpersonen bei der Entwicklung der Bewertungsinstrumente eine relevante Einflussgröße zugeschrieben wird. Gleichwohl ist im

Rahmen dieser Arbeit die Rolle des Schulqualitätsrahmens eher gering einzuschätzen. Dies hängt unter anderem damit zusammen, dass der Schulqualitätsrahmen in Hinblick auf Messinstrumente zu Inklusion von den Interviewpersonen nicht mehr genannt wird, d.h. wenn es um konkrete Instrumente bezüglich Inklusion geht, gibt es keinen Verweis auf den Schulqualitätsrahmen, sondern nur, wenn es um das Instrumentenset im Allgemeinen geht. Dies lässt sich in vielen Fällen damit erklären, dass Inklusion in fast allen Schulqualitätsrahmen nicht explizit genannt wird. Des Weiteren lässt sich der Verweis auf den Schulqualitätsrahmen zum Teil auch so deuten, dass dies auch als Strategie genutzt wird, um das eigene Handeln nach außen zu legitimieren.

Bewertungsinstrumente aus dem ersten Zyklus In den Interviews ging es bis auf wenige Ausnahmen um die Bewertungsinstrumente aus dem zweiten Zyklus der Schulinspektionen, also um die Entwicklung der Instrumente, die Anwendung fanden, als die Schulen zum zweiten Mal evaluiert wurden. Dies ist von Relevanz, da fast alle betrachteten Schulinspektionen zwischen 2004 und 2006 eingeführt wurden (→ Kapitel 3). Der erste Zyklus, d.h. die Durchführung der ersten Inspektion bei allen Schulen, dauerte ca. fünf Jahre. Die meisten Bundesländer nahmen das Ende des ersten Zyklus‘ zum Anlass, um die eigenen Bewertungsinstrumente anzupassen bzw. zu verändern. Zu der Zeit, als die BRK von Deutschland ratifiziert wurde, waren also in allen Bundesländern (außer Schleswig-Holstein) bereits Bewertungsinstrumente aus dem ersten Zyklus vorhanden.

Die Bundesländer unterscheiden sich darin, ob die bestehenden Messinstrumente leicht angepasst wurden oder ob tatsächlich größere Veränderungen stattgefunden haben, die sich unter anderem in einer neuen Struktur widerspiegeln. *„Ähm also es war eigentlich immer klar oder das wurde auch immer betont, dass man nicht sozusagen den alten, das alte [anonymisiert, ap] in das Neue überführen sollte, sondern dass man sich davon eigentlich in dieser Sichtweise ähm lösen musste“* (Int. 7_3). Dennoch merkt die Interviewperson an, dass der Versuch einer Neuausrichtung in der Praxis nicht geklappt hat (Int. 7_4). So haben die Bewertungsinstrumente aus dem ersten Zyklus auch die Entwicklung von Bewertungsinstrumenten für den zweiten Zyklus mitgeprägt, z. B. indem sie sich an den vorhandenen Instrumenten oder Strukturen orientiert haben.

In Bezug auf Inklusion fand bei vielen Schulinspektionen ein Abgleich davon statt, inwiefern die Bewertungsinstrumente aus dem ersten Zyklus bereits das Konzept als solches abdecken (Int. 12_26; 14_38). Es gab unterschiedliche Einschätzungen darüber, ob Inklusion bereits durch die bisherigen Instrumente abgedeckt ist. Dieser Abgleich führte zu unterschiedlichen Reaktionen:

- Die eigene Deutung von Inklusion findet sich in den Instrumenten wieder: Demnach wurden auch keine zusätzlichen Instrumente entwickelt. *„In der zweiten Runde ham wir dann zum einen, äh, noch mal in einer längeren Phase tatsächlich überprüft, ob wir mit den vorhandenen Bewertungskriterien Inklusion abbilden können, oder ob wir zusätzliche Kriterien kreieren müssen. (.) Und haben eigentlich, je länger wir uns- Äh, es- es gab dazu eine kleine Arbeitsgruppe aus Förderschullehrern und äh Regelschullehrern. Und je mehr beschäftigt haben, desto mehr haben wir eigentlich gemerkt: Nein, wir wollen ganz bewusst da keine weiteren Kriterien einführen, weil eigentlich, wenn ein- ein Umgang mit Vielfalt gelingt, dann is das guter Unterricht“* (Int. 15_3).
- Die eigene Deutung von Inklusion findet sich in den Instrumenten wieder, aber nicht unter dem Begriff „Inklusion“. Des Weiteren gibt es das Szenario, dass die eigene Herangehensweise an Inklusion sich in den Instrumenten widerspiegelt, aber nicht explizit unter dem Begriff „Inklusion“ gefasst wird. Hier wurden zum einen bestehende Instrumente als „inklusionsori-

- entiert“ gedeutet oder es wurden zum anderen aus strategischen Gründen keine Änderungen vorgenommen, z. B. um weiterhin kohärent zu den Setzungen des Ministeriums zu bleiben.
- Die eigene Deutung von Inklusion findet sich in den Instrumenten kaum wieder: In diesem Fall gibt es zwei Reaktionen. Zum einen die Reaktion, dass neue Instrumente zu Inklusion entwickelt wurden. Aber zum anderen gibt es auch den Fall, dass eine Interviewperson konkrete Vorstellungen davon hat, wie – ausgehend von der eigenen Deutung – Inklusion in den Instrumenten berücksichtigt werden könnte, aber die Person konnte sich innerhalb der eigenen Schulinspektion nicht durchsetzen (nähere Ausführung folgt).

Zeitliche sowie finanzielle Bedingungen Die zeitlichen Rahmenbedingungen scheinen eine untergeordnete Rolle gespielt zu haben. So wird lediglich in zwei Fällen berichtet, dass sie nur wenig Zeit hatten, um Instrumente zu Inklusion zu entwickeln (Int. 17_30; 18_2; 10_30): *„Also nochmal die Vorgehensweise für Bundesland [anonymisiert] war insofern ungewöhnlich, als dass sie im pragmatischen Sinne innerhalb eines Jahres - das war die Vorgabe der Ministerin - eigentlich ein Verfahren auf die Beine stellen, dass wir ein Verfahren auf die Beine stellen mussten - das Akzeptanz führt, das anerkannt wird und das auch ne Grundlage für ein neues Verfahren zur externen Evaluation darstellt“* (Int. 18_2). Dies habe zur Folge, dass die entwickelten Bewertungsinstrumente kaum erprobt wurden. Eine Interviewperson sieht dies als nicht dramatisch an, weil die Schulinspektionen es sowieso nie allen recht machen könnten (Int. 17_30). Dahingehend weist eine andere Interviewperson darauf hin, dass die fehlende Zeit dazu geführt hat, dass man sich stark an bestehenden Instrumenten aus anderen Bundesländern orientiert hätte (Int. 18_2) und die eigenen Bewertungsinstrumente in Bezug auf Niveau und Bandbreite mit anderen Kriterien nur begrenzt mithalten könnten (Int. 18_32). (Diese Aussagen beziehen sich jedoch auf das allgemeine Instrumentenset und nicht auf die Instrumente zu Inklusion.) In Bezug auf die Entwicklung von Bewertungsinstrumenten zu Inklusion erwähnt eine Person, dass die Auseinandersetzung mit Inklusion begann, als die Entwicklung von Instrumenten eigentlich schon abgeschlossen war (Int. 10_31) und sie somit nur begrenzt zeitliche Kapazitäten zur Verfügung hatten. Es wird jedoch nicht thematisiert, inwiefern sich dies auf die konkrete Gestaltung der Instrumente ausgewirkt hat.

Ähnlich wie bei den zeitlichen Rahmenbedingungen scheinen die finanziellen Bedingungen kaum eine Rolle zu spielen. So gibt es zwar bei mehreren Schulinspektionen einen Sparauftrag. Dies führt dazu, dass weniger Personal vorhanden ist (Int. 2_11) und kostengünstigere Messformate entwickelt werden (Int. 12_33). Aber es gab keine Aussagen darüber, dass es sich – außer auf die Anzahl der Indikatoren – inhaltlich auf die Auswahl der Bewertungsinstrumente ausgewirkt hat. Dennoch kann auch eine begrenzte Anzahl an Indikatoren sich insofern auf die Auswahl an Instrumenten auswirken, dass nur Teilaspekte von Inklusion abgefragt werden können.

Funktionen von Schulinspektionen Wie bereits in Kapitel 3 dargestellt, werden Schulinspektionen nach Landwehr (2011) vier Funktionen zugeschrieben: Kontroll- und Impulsfunktion, Normendurchsetzung sowie Wissensermittlung.

Nach den Aussagen der Interviewpersonen haben die Schulinspektionen zwar verschiedene Funktionen zu erfüllen, aber es werden unterschiedliche Positionen dazu besetzt, was als primäre Funktion(en) angesehen wird. Auch gibt es innerhalb der Schulinspektionen Diskussionen darüber, wie man sich in Bezug auf die einzelnen Funktionen positionieren sollte (z. B. Int. 6_7; 16_20): *„was ist unsere Aufgabe als Schulinspektion [anonymisiert] und was ist die Aufgabe*

der schulfachlichen Aufsicht. [...] Das führt immer wieder zu neuen Diskussionen. [...] Das ist ne Grundsatzdebatte, die noch nicht zu Ende geführt ist“ (Int. 6_7). In Abhängigkeit vom Verfahrensschritt oder von der jeweiligen Schule kann variieren, was als relevante Funktion wahrgenommen wird: *„Ja manchmal is sie ja- s-es-es gibt manchmal- äh-hm-hm ploppen von den-von den vier äh Funktionen ploppen manchmal einer mehr auf, manchmal- also dis wandelt sich auch von Schule zu Schule auch“* (Int. 14_21; siehe auch 13_34). Falls eine Schule beispielsweise gut aufgestellt ist, geht es eher nicht darum zu gucken, ob bestimmte Standards erfüllt sind (-Kontrollfunktion), sondern eher um die Frage, wo die Schule sich noch weiterentwickeln kann (-Impulsfunktion). *„Schulinspektion hat irgendwie vier Funktionen und die sind miteinander nich gerade besonders gut kompatibel“* (Int. 13_31). Im Folgenden wird nun dargestellt, wie sich die Befragten zu möglichen Funktionen von Schulinspektionen positionieren und welche Implikationen sich aus den einzelnen Funktionen ergeben: In den Interviews taucht als eine mögliche Funktion auf, ob die Schulinspektionen verantwortlich dafür sind, **dass bestimmte Gesetze an den Schulen eingehalten werden**. Hier zeigt sich, dass viele Interviewpartner sich von dieser Funktion explizit distanzieren (Int. 13_25; 12_18) bzw. diese nicht benennen. So sollen die Bewertungsinstrumente der Schulinspektion zwar – wie bereits oben erwähnt – mit den rechtlichen Vorgaben kohärent sein, aber die Kontrolle davon sei nicht die Funktion der Schulinspektionen, insbesondere nicht, weil dies als Aufgabe der Schulaufsicht gesehen wird. So gibt es nur ein Bundesland, wo der Rechtsanspruch auf Inklusion überprüft wird. Dennoch verweisen die Instrumenteentwickler_innen häufig auf rechtliche Vorgaben, um die Instrumente zu legitimieren. Des Weiteren haben rechtliche Vorgaben wie die BRK die Entwicklung von Instrumenten initiiert. Wie bereits oben dargestellt, sind sie eher als Handlungsrahmen zu verstehen, aber die Überprüfung dessen, ob an Schulen Recht eingehalten wird, steht bei den meisten Schulinspektionen nicht im Mittelpunkt.

Als weitere Funktion von Inspektionen wird **Bildungsmonitoring** genannt (Int. 14_19; 11_6; 17_53; 5_26; 13_17; 16_25). Demnach geht es darum, dass die Daten an Schulträger bzw. in aggregierter Form für den Bildungsbericht auf Bundesländerebene weitergegeben werden. Dies bietet für die Kultusministerien als Rezipienten der Daten die Möglichkeit einen Überblick über die Rahmenbedingungen, aber auch über die tatsächlichen Prozesse an den jeweiligen Schulen zu erhalten. Dies hat zur Folge, dass in den Bewertungsinstrumenten Aspekte berücksichtigt werden, worüber die Rezipienten gerne mehr Informationen hätten bzw. kann es auch sein, dass die Schulinspektionen selbst antizipieren, welche Daten aufgrund der bildungspolitischen Schwerpunktsetzung von Relevanz sein könnten (Int. 13_17). Die Funktion des Bildungsmonitorings impliziert jedoch nicht, dass es Vergleichsmöglichkeiten zwischen Schulen geben soll. Demnach ist es in den Bundesländern – bei den obligatorischen Instrumenten – technisch möglich, die Schulen zu vergleichen, und dies wird mindestens in einem Fall nach Angaben eines Instrumenteentwicklers auch für die Behörden gemacht: *„Also wir aggregieren die [Daten, ap] und wir geben die raus, aber nur an die Schulaufsicht und an die staatlichen Behörden. Das ist schade, weil ich denke - oder wir denken - erstes Mal denken wir, dass ein Teil unserer Arbeit nicht transparent ist. Wir müssen es machen. Zum anderen ist es so, dass der Vergleich mit anderen Schulen durchaus was hergibt - ich verstehe es auch nicht ganz, warum man da nicht weitergeht, weil ein Vergleich einer einzelnen Schule mit einem aggregierten Wert der Schulart in einem bestimmten Bezirk oder was. Ja kein Ranking ist. Also sagen wir mal als Orientierungsrahmen für die Schulen und wir legen großen Wert drauf, den Schularten zu zeigen, also der Schulaufsicht. Also wir spiegeln z. B. keine Mittelwerte zurück, sondern wir spiegeln Verteilungen zurück. Man kann sogar - nicht*

schlecht - bei schlecht bewerteten Kriterien sehen, dass es da Schulen gibt, die da immer noch gut abschneiden. Um das geht es uns“ (Int. 17_54).

Des Weiteren wird den Schulinspektionen auch von Seiten der Interviewpartner **eine normendurchsetzende Funktion** (Int. 17_23; 9_4; 14_22; 13_31; 16_32) zugeschrieben, denn sie legen fest, welche Kriterien von guter Schule als verbindliche Orientierung gelten sollen und versuchen, diese durch die Evaluation durchzusetzen: *„Also diese Orientierungsfunktion - also nur so als Beispiel. In Hessen gab es ja den Arbeitskreis Qualität an Schulen über 20-30 Jahre - was weiß ich - ich habe mal bei einer Veranstaltung in Hessen - da war so - die haben mal so ne Bestandsaufnahme, mal so ne große Tagung gemacht und da war der Staatssekretär und da hat er gesagt hat - wenn die Evaluation etwas gebracht hat, dann, dass wir heute wissen, was Qualität sein soll“ (Int. 17_23).* Durch ihre Kriterien sensibilisieren sie die Schulen für bestimmte Aspekte und machen beispielsweise auf Pflichten aufmerksam. Durch die Vorbereitung auf die anstehende Inspektion werden die Schulen auf die Anforderungen aufmerksam bzw. im Nachhinein dafür sensibilisiert, wenn sie von den Schulinspektionen Rückmeldungen erhalten. In diesem Zusammenhang stellt sich die Frage, inwiefern sich die Instrumenteentwickler_innen nicht nur als normendurchsetzend, sondern auch als normsetzend wahrnehmen. Danach würden sie selbst als Policymaker agieren und nicht nur vorgegebene Normen operationalisieren und messen. Hier zeigt sich ein gemischtes Bild: Zum einen wird die Auswahl an Instrumenten – wie oben bereits dargestellt durch den Verweis auf die Bildungsforschung u.ä. begründet. Danach entsteht der Eindruck, dass die Instrumenteentwickler_innen sich selbst eher in einer passiven Rolle sehen, da sie die von anderen festgelegten Normen von guter Schule in Indikatoren übersetzen, messen und damit versuchen, diese Normen durchzusetzen. Auf der anderen Seite gibt es auch Hinweise darauf, dass sich die Instrumenteentwickler_innen durchaus auch als normsetzend wahrnehmen (Int. 15_15; 9_44; 1_43): *„Aber ich glaube schon, dass wir natürlich auch Standards setzen und im Grunde - das haben wir jetzt zum Beispiel in dem Bereich gemacht, indem wir gesagt haben, die Wissenschaft ist sich da so dermaßen einig, dass eben diese genaue Passung zum Lernfortschritt ganz ganz wichtig ist, um da maximal große Lernfortschritte zu erreichen und um auch eben dieses gesamte Spektrum, also nicht nur die allerschwächsten, sondern auch die allerstärksten entsprechend zu fördern. Da denke ich - da haben wir schon so nen bißchen eine Richtung vorgegeben“ (Int. 1_43).* In Bezug auf Inklusion wird die normensetzende Funktion deutlich, wenn ein Interviewpartner es als Aufgabe der Schulinspektion ansieht, die Auseinandersetzung mit Inklusion im eigenen Bundesland mit zu prägen. Dabei geht es ihm darum, Inklusion stärker in den Instrumenten zu verankern und auf eine bestimmte Art und Weise zu definieren: *„Also dann - wird es schon - wie sage ich mal meine Intention sein, hier im Fachbereich darauf zu drängen, dass wir jetzt also dann dieses Wort [Inklusion, ap] dann doch auch mal benutzen. Im Sinne von <wir müssen auch unseren Beitrag leisten ein anderes Inklusionsverständnis zu entwickeln>“ (Int. 9_44).*

In Anknüpfung an die norm(durch)setzende Funktion, wird Schulinspektionen auch eine **Rechenschaftslegungsfunktion bzw. Kontrollfunktion** zugeschrieben. Demnach hat die Schulinspektion die Aufgabe zu überprüfen, inwiefern bestimmte Standards an Schulen erfüllt sind. Hier zeigt sich in Deutschland die Tendenz, dass sich viele Schulinspektionen seit ihrer Einführung von dieser Funktion immer mehr distanzieren. So gibt es nur einige wenige interviewpartner, die der Kontrollfunktion weiterhin eine wichtige Rolle beimessen (Int. 8_44; 13_31). Ansonsten hat die Kontrollfunktion für die Schulinspektionen an Relevanz verloren (Int. 4_21; 6_21; 16_1; 10_8; 17_12) bzw. spielt überhaupt keine Rolle mehr (Int. 9_30; 18_13; 5_37). Es gibt diverse Gründe, die diese Entwicklung erklären können: Beispielsweise das Bedürfnis, sich noch stärker von der Schulaufsicht abzugrenzen und ihnen somit die Kontrollfunktion

zuzuschreiben (Int. 9_23). Des Weiteren vermuten einzelne Instrumenteentwickler_innen, dass die Kontrollfunktion dazu führt, dass die Schulen so genanntes „Window Dressing“ betreiben, d.h. die Schulen bemühen sich darum, sich möglichst gut während der Inspektion darzustellen, ihre Stärken zu betonen und ihre Schwächen zu verstecken (Int. 9_21; 4_22). In dem Fall fehlt es den Schulinspektionen an Informationen darüber, welche Schwächen die Schulen tatsächlich haben und vor welchen Herausforderungen sie stehen. Zusätzlich wird die frühere Dominanz der Kontrollfunktion als Ursache dafür gesehen, dass die Schulinspektionen von Seiten der Schulen eine geringe Akzeptanz erfahren. Durch die stärkere Ausrichtung an der Impulsfunktion erhoffen sich die Schulinspektionen daher mehr Akzeptanz von Seiten der Schulen. Auch wenn die Relevanz der Kontrollfunktion abnimmt, sind sich die Interviewpersonen dessen bewusst, dass sie von den Schulen häufig weiterhin als Kontrollinstanz wahrgenommen werden (Int. 4_22; 11_16; 16_19; 9_18): *„Ähm wir kommen nicht mehr daher mit dem Gestus wir beurteilen euch und wissen wie es ist, sondern wir haben den Begriff Deutungsangebot geprägt. Ähm das wir auf der Grundlage der, mit der Bewertungsgrundlage, die wir als Folie immer im Hinterkopf haben, also den Orientierungsrahmen und die anderen Dinge. Ein Blick auf Schule, ein sehr tiefen Blick auf Schule werfen und vor diesem Hintergrund der Schule ein Deutungsangebot ihrer Qualität machen können. Es gibt aber auch andere Folien, die zum Beispiel Schule im Kopf haben kann“* (Int. 16_19). Demnach wird von Seiten der Schulinspektionen und den Schulen unterschiedlich wahrgenommen, welche Funktion die Schulinspektionen übernehmen. Die Perspektive der Schulen lässt sich dadurch erklären, dass die Teilnahme an der Schulinspektion in allen Fällen (außer Schleswig-Holstein) verpflichtend ist und die Ergebnisse nicht nur der Schule, sondern auch der Schulaufsicht und anderen Akteuren kommuniziert werden bzw. in manchen Fällen diese auch online zugänglich sind (→ Kapitel 3).

Während die Kontrollfunktion aus der Perspektive der Schulen sehr präsent ist, wird von den meisten Entwickler_innen die **Impulsfunktion für Schulentwicklungsprozesse** als primäre Aufgabe der Schulinspektionen angesehen (Int. 8_43; 9_26; 12_9; 14_19; 18_7; 4_21; 5_26; 6_8; 7_11; 13_31; 16_19; 17_12; 11_16): *„Und der vierte Aspekt, der in der Schulinspektion, glaube ich, in den letzten Jahren den höchsten – die höchste Gewichtung erfahren hat, ist der Impuls zur Schulentwicklung zu gehen. Also mit Daten gestützten Rückmeldungen Schule dazu befähigen zu erkennen: Was läuft gut, aber was läuft noch nicht gut, woran müssen wir arbeiten, und wo können wir ansetzen zu arbeiten“* (Int. 13_31). Die stärkere Hinwendung zur Impulsfunktion lässt sich unter anderem dadurch erklären, dass die Schulinspektionen von Seiten der Schulen die Rückmeldung erhalten haben, dass diese sich nicht in ihren schulspezifischen Rahmenbedingungen abgeholt fühlen, wenn von allen Schulen die Erfüllung der gleichen Standards erwartet werden (Int. 5_23). Möglicherweise lässt sich die Hinwendung zur Impulsfunktion in manchen Bundesländern auch durch die Rollenidentität der Inspektor_innen erklären: Wie bereits oben erwähnt sind viele Inspektor_innen (ehemalige) Lehrkräfte bzw. ehemalige Mitarbeiter_innen der Schulaufsicht. In dem Zusammenhang erzählte eine Inspektorin in einem informellen Gespräch über die stärkere Hinwendung zur Impulsfunktion, dass sich damit nun die Inspektor_innen wohler fühlen würden, weil dies eine andere Form der Interaktion gegenüber Lehrkräften ermöglicht und diese Funktion der Tatsache gerecht wird, dass sie von schulischen Vertreter_innen als (ehemalige) Kolleg_innen bzw. als critical friend wahrgenommen werden. Dagegen gab es im Rahmen eines anderen informellen Gesprächs die Hintergrundinformation, dass Inspektor_innen – in diesem Fall ehemalige Schulaufsichtsmitarbeiter_innen – die Kontrollfunktion in ihrem Handeln noch stark verinnerlicht hätten, obwohl dies von Seiten

der Schulinspektionsleitung nicht mehr als primäre Funktion der jeweiligen Schulinspektion angesehen wurde.

Durch den Fokus auf die Impulsfunktion erhoffen sich die Interviewpersonen mehr Akzeptanz von Seiten der Schule (z. B. Int. 10_8; 18_11) sowie mehr Wirksamkeit in Bezug auf schulische Entwicklungsprozesse: *„Und wir wollten gern, wir wollten gern ein Verfahren so haben, dass zumindest die Wahrscheinlichkeit, dass man ein bisschen länger damit arbeitet, weil mans ja auch für sinnvoll hält. Nachweise kann man da nicht führen, glaube ich“* (Int. 10_8; siehe auch 18_11). Jedoch besteht bei einzelnen Instrumenteentwickler_innen Skepsis darüber, inwiefern mit der eher sanfteren Herangehensweise, die mit der Fokussierung auf die Impulsfunktion einhergeht, auch wirklich Schulentwicklungsprozesse angestoßen werden: *„Eigentlich ist die – die wirklich erfolgreichste Funktion von Schulinspektion ist: Rechenschaftslegung. Das ist vielleicht nicht die Wunsch erfolgreichste Funktion. Wie gesagt, mir wärs lieber, wir könnten sagen, - es-es gelingt uns Schulentwicklung anzustupsen“* (Int. 15). Der Fokus auf die Impulsfunktion wird auch als *„stumpfes Schwert“* (Int. 9_27) beschrieben. Demnach geht mit dieser Funktion zwar die Hoffnung einher, dass schulische Entwicklungsprozesse dadurch stärker vorangetrieben werden, aber einzelne Erfahrungen aus der Praxis zeichnen ein anderes Bild. Diese Einschätzung wird durch die Studie von Würster et al. (2016) geteilt, da sanftere Verfahren - trotz tendenziell höherer Akzeptanz – nicht zu mehr Aktivitäten von Seiten der Schulleitung führen. Wie bereits dargestellt, argumentiert Landwehr, dass die Kontrollfunktion – im Vergleich zur Impulsfunktion – *„die wichtigere Existenzberechtigung“* (Landwehr 2011, S. 58) für die Schulinspektion darstellt: Dies macht er daran fest, dass in Bezug auf Schulentwicklung die Impulsfunktion im Vergleich zur internen Evaluation von Seiten der Schule möglicherweise weniger wirksam ist. Jedoch die Kontroll- bzw. Rechenschaftslegungsfunktion *„einem differenzierten Qualitätsverständnis und den Glaubwürdigkeitsansprüchen einer aufgeklärten Gesellschaft“* (ebd., S. 58) entspricht. Demnach könnte die stärkere Orientierung an der Impulsfunktion und Distanzierung von der Kontroll- bzw. Rechenschaftslegungsfunktion dazu führen, dass das Alleinstellungsmerkmal von Schulinspektionen verschwindet.

Die Impulsfunktion kann verschiedene Implikationen bezüglich der Entwicklung von Instrumenten mit sich ziehen. So liegt es nahe, dass nur diejenigen Aspekte in den Bewertungsinstrumenten aufgeführt werden, bei denen bei den Schulen noch Handlungsbedarf gesehen wird (Int. 2_13; 7_13; 4_24) bzw. es wird versucht, eine Balance zwischen bereits funktionierenden und verbesserungswürdigen Aspekten zu halten. So sollten aus Motivationsgründen auch Aspekte aufgeführt werden, die schon gut laufen (z. B. Int. 10_26). Des Weiteren ergibt sich aus der Impulsfunktion, dass den Schulen mehr Mitbestimmungsrecht eingeräumt wird (Int. 9_6; 4_21; 6_9). Dies kann sich beispielsweise darin äußern, dass es einen Pflicht- und einen Wahlbereich gibt: *„Die Beteiligung der Schulen war eigentlich nen Reflex darauf, dass viele Schulen gesagt haben <Okay, wir haben euch nicht gerufen, wir wollen euch eigentlich nicht. Wir sind eigentlich nicht gefragt worden. Und ihr seid so über uns hergefallen und naja und dann wart ihr ne Woche da und dann wart ihr wieder weg.> Und äh. Und die Schulen größer zu beteiligen - ist jetzt nen Versuch von uns - sage ich mal - nen Stück weit Beziehung herzustellen im Vorfeld, um dann den Schulbesuch nen Stück weit auch gemeinsam zu machen.“* (Int. 9_18). Darüber hinaus spiegelt sich die Möglichkeit zur Mitbestimmung auch darin wider, dass einige Schulinspektionen mittlerweile zwischen Deskription und Evaluation unterscheiden. Die Idee dahinter ist, dass die Schulinspektionen nur noch Daten zurückmelden und es den Schulen überlassen wird, wie diese zu bewerten sind (nähere Ausführung folgt).

Die Darstellung der einzelnen Funktionen hat bereits erste Hinweise gegeben, inwiefern die Kombination von verschiedenen Funktionen problematisch sein kann: *„Und das ist das Problem, das wir haben, wenn wir diese beiden Funktionen [Kontroll- und Impulsfunktion, ap]) vermischen. Dann haben wir auf der einen Seite die Rechenschaftslogik, die die Logik erfordert: Zeige mir, wie gut du bist. Zeige mir deine positiven Seiten, damit du gut abschneiden kannst – im Sinne von Rechenschaft. Gleichzeitig ist aber die Impulsgebungsfunktion so angelegt, dass man auch die Schwächen der Schule identifizieren möchte, um der Schule dann sagen zu können – hier und da wirkt es so, als wenn du noch nicht so gut aufgestellt bist, - ähm – da könntest du z. B. gucken, ob es in deinem Sinne ist, dass du dort noch etwas machen möchtest“* (Int. 4_22; 13_32).

Das Spannungsfeld zwischen Kontroll- und Impulsfunktion spiegelt sich auch in der Formulierung von Instrumenten wider. Aus der Perspektive der Kontrollfunktion liegt es nahe Standards zu formulieren, z. B. Mindeststandards, die alle Schulen erfüllen sollen. Dann wäre es möglich, dass alle Schulen ähnlich abschneiden, weil alle die an sie formulierten Standards erfüllen. Dies ist beispielsweise bei der Bund-Länder-Inspektion der Fall, da beispielsweise im Jahr 2009 alle deutschen Schulen im Ausland das Gütesiegel „Exzellente Deutsche Auslandsschule“ erhalten haben (Bundesverwaltungsamt 2010, S. 10). Ausgehend von der Impulsfunktion wird erfasst, welche Prozesse an der Schule laufen und welche Prozesse umgesetzt werden müssen, um bestehende Probleme und Herausforderungen zu meistern. Es würde keine sachliche Bezugsnorm gelten, sondern es könnte eine einzelschulspezifische Bezugsnorm angelegt werden. Demnach würde es um die Frage gehen, wohin sich die Schulen unter Berücksichtigung ihrer aktuellen Situation entwickeln sollten. Demnach können durch die Kontrollfunktion bildungspolitische Interessen berücksichtigt werden: *„Wir sagen einerseits wollen wir uns schulbezogen, schulspezifisch da standardspezifisch ausrichten. Und wenn ich Schule ernst nehmen müssen wir auf Verpflichtungen ganz verzichten. Eigentlich auch auf Bewertung. Und andererseits gibt es aber eine Erwartung von der Politik bestimmte Sachen auf jeden Fall in den Blick zu nehmen“* (Int. 5_22). Die Impulsfunktion kann dagegen implizieren, dass das „Wohin“ in Bezug auf schulische Entwicklungsprozesse vorab nicht festgelegt wird, sondern die Schulinspektion dies gemeinsam mit den Schulen vereinbart. Demgemäß würde in Interaktion mit den Schulen erarbeitet werden, was gute Schule für die jeweilige Schule bedeutet.

Zusammenfassend zeigt sich, dass sich die Mitarbeiter_innen der Inspektion für mehrere Funktionen verantwortlich fühlen, z. B. Bildungsmonitoring. Dabei zeichnet sich die Tendenz ab, dass die Kontrollfunktion an Relevanz verliert, während die Impulsfunktion an Bedeutung zunimmt. Hierbei ist jedoch zu berücksichtigen, dass sich die Interviewpersonen darin unterscheiden, welche Relevanz der Kontrollfunktion weiterhin zugeschrieben wird. In Bezug auf Inklusion spiegelt sich die Festlegung auf die primäre Funktion unter anderen darin wider, welche Erwartungen an Schulen formuliert werden und ob die Instrumente zu Inklusion sich im Wahl- oder Pflichtbereich befinden.

Empirische Messung von Schulqualität Die Schulinspektionen stehen nicht nur vor der Herausforderung festzulegen, was eine gute Schule ist, sondern es geht auch um die Frage, wie sich dies empirisch überprüfen lässt. In den Interviews wird eine Vielzahl von Anforderungen genannt, die im wissenschaftlichen Bereich Gütekriterien darstellen: Die Kriterien sollen insofern valide sein, als dass die Instrumente das Konstrukt messen sollen, was abgebildet werden soll (Int. 14_25; 13_6). Darüber hinaus besteht auch die Sorge, dass sozial erwünschte Antworten erfasst werden: *„So dass wir dann nochmal zusätzliche Itembatterien drin hatten im Fragebogen. Wo ich aus heutiger Sicht sagen würde, so richtig hilfreich warn die auch nich. Weil*

man grade in diesem Bereich Inklusion so ganz schnell in so soziale Erwünschtheit-Bereiche kommt, die ganz schnell Deckeneffekte erzielen. Also da finde ich, wenn ich mal ganz offen und ehrlich bin, haben wir da kein Glanzstück (.) empirischer Art hingelegt“ (Int. 13_6).

Zusätzlich sollen die Instrumente reliabel sein (Int. 2_16; 7_7; 10_28) und die Konstrukte sollen trennscharf erhoben werden (Int. 17_29; 10_28). Darüber hinaus wird als Anspruch formuliert, dass ein Kriterium beobachtbar sein muss bzw. es muss mit den Methoden der Schulinspektionen gemessen werden können (Int. 1_18; 2_24; 14_24; 8_41; 6_4; 5_4; 7_6). So gibt es Aspekte, die von Seiten der Interviewpartner in Bezug auf Schulqualität maßgeblich sind, aber nicht durch die Schulinspektionen gemessen werden können. Dies kann beispielsweise daran liegen, dass die Daten nicht vorliegen bzw. die Messung der Daten durch die Schulinspektionen nicht möglich ist (Int. 13_22; 15_5; 10_14). So argumentiert eine Interviewperson, dass manche Ziele, z. B. Befähigung zum guten Leben, erst nach Abschluss der Schulzeit gemessen werden können und es von Seiten der Schulinspektion *„keine weitere Verfolgung des Lebensweges gibt“* (Int. 10_14). Der Aspekt der Messbarkeit wirkt sich darauf aus, welches Verständnis von guter Schule sich in den Instrumenten widerspiegeln kann *„Also die, es geht eigentlich in Schule darum, in dieser Reihenfolge, meine ich, dass haben andere aber auch schon gesagt, die Menschen zu stärken und die Dinge zu klären. Beides. Und manchmal hat man den Eindruck es geht in der aktuellen Zeit viel darum Messbares in den Mittelpunkt zu rücken und das missachtet eigentlich den anderen Strang [...] des schulischen Auftrags. Ähm also es ist nicht alles messbar und nicht Messbares kann trotzdem sehr bedeutsam für das, was Schule zu leisten hat, sein“* (Int. 5_12). Demnach sind bestimmte Aspekte, die für die Instrumenteentwickler_innen in Bezug auf Schulqualität von Relevanz sind, empirisch nicht erfassbar: *„ja erfahren sie denn auch äh den, ich glaube er hat gesagt, den Spirito unserer Schule? Also zu sagen den Spirit, den Esprit, den Geist, man könnte vielleicht sogar fast sagen die Seele der Schule. Und ich denke nein. Und trotzdem ist das etwas, (...) was für eine Schule unglaublich wichtig ist. Was aber ganz schwer so ne Inspektion erfassen kann“* (Int. 8_41).

Des Weiteren berichtet eine Interviewperson von einem Item, das auf den gleichberechtigten Umgang zwischen Mädchen und Jungen abzielt: *„Ich hab das glaub ich ein einziges Mal mit nicht erfüllt bewertet. Dann ist die Frage, was macht die Schule damit. So äh die Einzelerückmeldung kann man natürlich nicht geben, ist ja die Lehrkraft identifizierbar und ich meine, dass zeigen dann auch die allgemeinen Auswertungen, dass dieses Item keine Aussagekraft mehr hat, weil es nicht prägnant genug ist und letztendlich das, was wir vielleicht, dass wäre das einzige Item was wir zu Heterogenität haben, ähm eigentlich andere Items zu Heterogenität geben müsste“²*. Demnach verlieren einzelne Konstrukte an Relevanz, wenn sie nur durch wenige Indikatoren bzw. Items abgedeckt werden.

Eine weitere Anforderung an die Messinstrumente ist, dass diese möglichst für alle Schulen anwendbar sein sollen. Dies impliziert, dass diese unabhängig von Schulart, Schulgröße sowie sozio-demographischen Bedingungen sind (Int. 16_18; 17_33; 15_18; 4_9). Dies ist jedoch herausforderungsvoll, da die Rahmenbedingungen und Prozesse zwischen einzelnen Schulen variieren: *„Und ich komme gerade aus einer Schule, wo Kollegen – kleine Schule- hervorragend zusammenarbeiten, also besser könnte man sich das nicht wünschen. Dieses strukturell aber gar nicht wirklich angelegt ist, es gibt eben keine Definition, kein wirkliches Controlling, kein Nachstellen zentral. So die werden ja diesen Bereich eher schwach erhalten. [...]. Ähm das sorgt aber, und das hat dann die Schulformexpertin gesagt, das sorgt mit Sicherheit für große Verwirrung, weil die Leute da sitzen, in der Präsentation hören, die Zusammenarbeit ist eher schwach und sich fragen, was*

² Aus Datenschutzgründen keine Nennung der Person.

ist das denn“ (Int. 16_18). Demnach zeigt sich bei der Entwicklung von Instrumenten die Herausforderung, dass die Instrumente für alle Schulen gelten sollen.

Der Aspekt der Messbarkeit spielt in Hinblick auf Inklusion insofern eine Rolle, da eine Vielzahl von Aspekten, z. B. Inklusion als Haltung, nur schwer durch die gängigen Instrumente gemessen werden kann. Die Messbarkeit bei der Identifikation von Schülergruppen spielt ebenfalls eine Rolle: So erklärt eine Interviewperson, dass Schüler_innen mit diagnostizierter Behinderung aufgrund der Etikettierung leichter identifiziert werden können (Int. 1_29). Zugleich wird der Trend zur Nicht-Etikettierung als Herausforderung gesehen, um die Umsetzung von Inklusion messen zu können (Int. 5_31). Zusammenfassend zeigen sich zwischen den Bundesländern Unterschiede darin, welche Relevanz der Einhaltung von wissenschaftlichen Standards beigemessen wird. Dies hängt unter anderem damit zusammen, dass der Anteil des wissenschaftlichen Personals in den einzelnen Schulinspektionen variiert.

Zusammenfassung der Kontextbedingungen Abschließend lässt sich festhalten, dass unterschiedliche Kontextbedingungen bei der Entwicklung von Instrumenten prägend wirkten. Dabei führten die jeweiligen Kontextbedingungen dazu, dass die Instrumenteentwickler_innen im Hinblick auf die Entwicklung von Indikatoren aktiv wurden. Des Weiteren zeigte sich, dass die Kontextbedingungen auf die inhaltliche Gestaltung der Indikatoren einwirkten und darauf, welche Erwartungen an Schule formuliert wurden.

In diesem Abschnitt wurde eine einzelfallübergreifende Darstellung gewählt, um das Spektrum an unterschiedlichen Rahmenbedingungen darstellen zu können. In Einzelfällen kann es jedoch variieren, welche Rahmenbedingungen eine Rolle gespielt haben bzw. wie diese eine Rolle gespielt haben. Dieser Aspekt wird in Abschnitt 8.3 näher dargestellt. Es wurde auch deutlich, dass die einzelnen Rahmenbedingungen miteinander interagieren. Auch auf diesen Aspekt wird in Kapitel 8.3 näher eingegangen.

8.1.4 Beispiel: Bewerten und Beschreiben in der Datenauswertung

Im Folgenden sollen mithilfe eines Beispiels verschiedene Aspekte des Sensemakings illustriert werden: Wie die Instrumenteentwickler_innen versuchen, zeitgleich verschiedenen Anforderungen, die sich aus den jeweiligen Rahmenbedingungen ergeben, gerecht zu werden und inwiefern ihnen dies gelingt. Des Weiteren wird anhand dieses Beispiels deutlich, dass es nicht nur von Relevanz ist, welche Daten erhoben werden, sondern auch wie diese ausgewertet und dargestellt werden.

Hierzu bietet sich ein Phänomen an, welches bei mehreren, aber nicht bei allen Inspektionen zu beobachten ist: der Verzicht auf Bewertungen und der reinen Deskription von Beobachtungen im einzelnen Abschnitten des Inspektionsberichts. Dies ist beispielsweise der Fall, wenn im Inspektionsbericht die Aussage getroffen wird, dass ca. 40 Prozent des beobachteten Unterrichts pünktlich anfängt oder dass im beobachteten Unterricht 80 Prozent der Zeit Frontalunterricht stattgefunden hat. Es bleibt nach Angaben der Instrumenteentwickler_innen bei dieser Rückmeldung und es wird keine Einschätzung darüber getroffen, ob dies gut oder schlecht sei. *„Also so in dem klassischen Sinne des Wortes „gut“, so schneidet bei uns keine Schule gut oder schlecht ab. Das hat etwas mit unserem Grundsatzverständnis zusammen, das ich versucht habe ihnen zu umreißen. Das wir also nicht bewerten, also das wir nicht sagen öh unsere Kriterien der guten Schule - seid ihr eine gute Schule oder seid ihr ne schlechte Schule, also ne failing school“* (Int. 9_49).

Im Gegensatz dazu wird es von den Instrumenteentwickler_innen als Bewertung verstanden, wenn Einschätzungen wie „gut“ oder „schlecht“ vorgenommen werden. In diesem Zusammenhang wird mehrfach auf die Ampel als Metapher verwiesen, d.h. bei „bewertenden“ Instrumenten würde die Schulinspektion deutliche Signale treffen, ob etwas beispielsweise im grünen, orangenen oder roten Bereich liegt. Es gibt verschiedene Begründungen für die „deskriptiven Abschnitte“ im Inspektionsbericht: So ist das Nicht-Beobachten von bestimmten Aspekten, z. B. Festhalten der Lernergebnisse **innerhalb einer beobachteten Unterrichtseinheit per se nicht schlecht**. Dennoch soll der Schule zurückgemeldet werden, wie häufig dies gesehen bzw. nicht gesehen wurde. Demnach möchte die Schulinspektion sich beispielsweise nicht positionieren, in wie viel Prozent der beobachteten Unterrichtseinheiten das Festhalten der Lernergebnisse stattfinden sollte: *„Und wenn man jetzt ein Unterrichtsbeobachtungsbogen solche Dinge einbaut und sagt, wir bauen nur die Dinge ein, die erwiesenermaßen immer stattfinden müssen. Im guten Unterricht muss immer gewährleistet werden, dass Schüler aktiv werden. So, dann bleibt man auf einer diffusen Ebene und man kann nur sehr undifferenziert die Dinge ähm in den Blick nehmen. Aber dadurch das wir sagen, es gibt kein Erwartungslevel ne [...],ap] und diese Interpretation der Ergebnisse muss in der Schule stattfinden“* (Int. 5_28).

Die deskriptive Darstellung bietet sich darüber hinaus dafür an, **den Schulen die Bewertung des Beobachteten zu überlassen**: *„dann weiß ich nicht, ob fünfzig gut oder schlecht ist. Das weiß nur die Schule - weil die Schule weiß, wo sie sich gerne sehen würde“* (Int. 4_18). Dahinter steht die Idee, dass die Schulen sich selbst überlegen sollen, welche Ziele sie verfolgen und die Schulinspektion nur als Datenlieferant fungiert: *„Also das wäre auch nur ne Rückmeldung an die Schule: <Ihr habt jetzt folgende Übergangsquoten - entspricht dies dem, was ihr eigentlich erreichen möchtet?>“* (Int. 4_32). Beziehungsweise wird es der Schule in Absprache mit der Schulaufsicht überlassen, wie sie die gelieferten Daten einschätzt. In Bezug auf die Rückmeldung von „deskriptiven“ Ergebnissen argumentiert ein Instrumenteentwickler: *„Das ist Ergebnis genug und da muss ich nicht noch draufsummeln und sagen jetzt, was ist jetzt gut oder schlecht. Das muss die Schule eigentlich selbst definieren bzw. wenn dann muss es eigentlich die Schulaufsicht oder die Schulverwaltung definieren - was für sie tolerabel ist und was nicht“* (Int. 17_34).

Die Nicht-Bewertung ermöglicht es auch, dass **die Situation der jeweiligen Schule berücksichtigt werden kann** (Int. 5_17): *„Wir sind nicht die, die Maßstäbe anlegen können, ob so etwas tolerierbar ist oder nicht“* (Int. 17_35). Demnach können bei der Einschätzung von Seiten der Schule bzw. Schulaufsicht diverse Rahmenbedingungen an der Schule mitgedacht werden, z. B. Ausstattung. Des Weiteren kann bei den Schulen als Bezugsnorm angelegt werden, wie sich die Schule in den letzten Jahren entwickelt hat: *„Aber wir sagen nicht, damit sind sie schon gut. Möglicherweise haben sie seit 5 Jahren Anstrengungen oder seit 3, um das zu erfüllen und sind bei 3 Prozent gestartet, sind jetzt bei 15.“* (Int. 5_17).

In den drei Begründungen spiegelt sich das Selbstverständnis der Schulinspektionen als Impulsgeber wider: *„Das is eben auch die Frage bei uns gewesen. sind wir mehr kontrollierend? (.) wertend? (.) Oder sind wir eher unterstützend, ähm (.) im Sinne von kooperativer äh ähm (.) Tätigkeit. Dass wir also versuchen, der Schule eine Außensicht zu geben, die von ihr genutzt wird, sich weiter zu entwickeln. Diese Hoffnung, äh-ähh, haben wir also verbunden mit der Schulinspektion [anonymisiert, ap]. Und dis is also wirklich die Frage gewesen, ja, also in welcher Rolle äh geben wir in die Schule hinein (.) und was wird hier äh betont. Dieser (2) konturierende, messende Aspekt oder der eher unterstützende und ähm (3). Das is im Ergebnis haben wir das unser Schulqualitätsrahmen [anonymisiert, ap], unsere Unterrichtsbeobachtung so ausgerichtet, dass wir eine Rückmeldung geben, so sehen wir das, so haben wir das erlebt. Ähm, geben auch Hinweise und Anregungen, machen jetzt*

aber keine dezidierte Wertung“ (Int. 12_10). Demnach möchten sie mit der Lieferung von Daten weniger kontrollieren, sondern die Schulen zur Reflektion anregen und dabei unterstützend wirken. Dies basiert auf der Annahme, dass Schulen eigenverantwortlich handeln und ihre Autonomie von Seiten der Schulinspektionen anzuerkennen ist. Es wird den Schulen bzw. der Schulaufsicht überlassen zu entscheiden, wo Handlungsbedarf besteht. In diesem Zusammenhang wird erneut deutlich, welche Rolle die schulischen Vertreter und die Schulinspektion spielen. Wie bereits oben erwähnt, sind diese Akteure in einzelnen Bundesländern nicht nur bei der Entwicklung der Instrumente beteiligt, sondern können bei der Evaluation an der eigenen Schule darüber entscheiden, wie die Daten zu bewerten sind. Die Schulinspektion sieht sich bei der Rückmeldung von Deskriptionen als Lieferant von neutralen Daten.

In diesem Zusammenhang ist jedoch kritisch zu hinterfragen, ob die Deskription von Daten bewertungsfrei sein kann. So impliziert die Erhebung von Daten bereits **eine Entscheidung darüber, welche Daten für die Messung von Schulqualität von Relevanz sind**. Hierbei bietet es sich an, auf den Prozess der Wertzuschreibung zu verweisen: Danach wird im Prozess zwischen Klassifizierung und Wertzuschreibung unterschieden. Eine Klassifizierung ist beispielsweise die Unterscheidung zwischen Berufsklassen und die anschließende Wertzuschreibung wäre die Hierarchisierung von Berufsklassen. Dabei würde die Klassifizierung dem „deskriptiven“ Vorgehen der Schulinspektionen entsprechen und der Schritt der Wertzuschreibung würde den Schulen überlassen werden. Wie bereits in Kapitel 4 dargestellt, ist auch die Klassifizierung nichts Naturgegebenes (vgl. Durkheim und Mauss 1987, S. 176). So werden durch Klassifizierungen bestimmte Perspektiven auf die Wirklichkeit sichtbar. In Bezug auf die Schulinspektion impliziert dies, dass Items wie „Die Lehrkraft ist Sprachvorbild“ insofern nicht mehr neutral sind, als hier bereits eine Entscheidung darüber getroffen wird, dass dieses Merkmal für die Messung von Schulqualität notwendig sei. Daran anknüpfend zeigt sich auch in der Praxis, dass **der deskriptive Teil von Seiten der Schule trotzdem als Wertung verstanden** wird: *„Aber das ist - eher die - ähm - die Selbstwahrnehmung: <Jetzt sind wir schlecht>. Also diese Form von <wir sind bewertet worden> - das ist eher die Wahrnehmung der Schule zu sagen <die haben uns jetzt schlecht bewertet>. Inspektoren sagen: <nein, wir haben hier keine Wertung ausgesprochen - wir stellen einfach fest - bei euch ist der Stand so und so. Das Bewerten - das passiert bei euch>“* (Int. 11_16). In diesem Zusammenhang stellt sich auch die Frage, inwiefern die Daten, die von Seiten der Schulinspektion erhoben werden, dem Verständnis von guter Schule von Seiten der jeweiligen Schule stets entsprechen. So könnte es sein, dass eine Schule andere Items formulieren würde, um die Qualität von Schule zu messen.

Durch die Erhebung von Daten werden nicht nur Schwerpunkte darin gelegt, welche Daten für die Messung von Schulqualität von Relevanz sind. **Darüber hinaus könnte man argumentieren, dass bei bestimmten Items der Erwartungshorizont hinsichtlich der Ergebnisse klar ist**. Beispielsweise gibt es das Item „Der Unterricht fängt pünktlich an“. Eine deskriptive Rückmeldung würde beispielsweise bedeuten, dass einer Schule zurückgemeldet wird, dass 30 Prozent des beobachteten Unterrichts pünktlich anfängt und keine Bewertung darüber stattfindet, ob dies zufriedenstellend ist. In diesem Fall stellt sich die Frage, inwiefern eine Bewertung durch die Inspektor_innen überhaupt noch notwendig ist, weil ein pünktlicher Unterrichtsbeginn im Fall von Deutschland eine allgemein anerkannte Norm darstellt. (Auch wenn es hier sicherlich Abweichungen dazu gibt, wie viel Prozent des Unterrichts pünktlich anfangen sollte.) **Des Weiteren wurde in den Interviews deutlich, dass die neutrale Position von Seiten der Schulinspektion nicht immer eingehalten wird**. Auf die Frage, wie damit umzugehen ist, wenn eine Schule für sich selbst in Ordnung findet, wenn die Lehrkräfte immer

unpünktlich sind, antwortet ein Instrumentenentwickler: *„Dann würden wir das so aufschreiben. Also dass aus unserer Sicht eine Nutzung der Unterrichtszeit nicht gewährleistet ist. Weder konzeptionell noch durch das Verhalten der Lehrkräfte oder der Leitung der Schule. Und wenn dann die Lehrkräfte zum Beispiel sagen, ist uns aber egal, dann ist über unsere Schleife auf der Ebene, auf der Basis des Berichts, in dem dann so etwas steht. Ne, wenn wir jeden Tag nach der großen Pause gehen ungefähr 10 Minuten verloren, das macht im Schuljahr ne 5700 Minuten pro Schüler. Das macht in der Schullaufbahn so und so viel Unterrichtsausfall aus. Das steht dann da. Das rechnen wir dann aus. Damit man das schön griffig machen kann. Und dann sitzen also auf der, mit dem Bericht, sitzen da Eltern und Schüler und lesen das und dann sagen Lehrer, ist uns aber egal. Und ne Schulaufsicht liest das auch. Denn die Schule muss hinterher auf der Basis von Entscheidungen, des Gremium Schulkonferenz mit Schülern, Eltern, Lehrkräfte an einem Tisch, muss dann halt die Schule, im Prinzip Schulleitung eine Zielvereinbarung schließen mit der Schulaufsicht. Und die Schulaufsicht liest auch den Bericht. Also wir melden Dinge zurück, auch Dinge die bedenklich sind, und das steht dann auch so etwas aus unserer Sicht ist dringend erforderlich sich mit dieser Organisationsform oder diesem Organisationsmodell äh für den Unterricht einer Schule zu befassen“ (Int. 5_29).*

Demnach wird insofern keine Wertung ausgesprochen, als nicht gesagt wird, dass die Unpünktlichkeit in Bezug auf Unterrichtsbeginn gut oder schlecht sei. Aber es werden im Bericht die Implikationen des Handelns dargestellt, die in diesem oben genannten Zitat eine deutliche Wertung enthalten. Die Wertung spiegelt sich nicht im Geben einer „Note“ wider, sondern in der Formulierung des Handlungsbedarfs. Wie bereits oben erwähnt, zeigt sich in der Anwendung von einem vermeintlich „deskriptiven“ Teil, dass sie sich primär in der Funktion des Impulsgebers und nicht in der Kontrollfunktion sehen. Hierbei ist jedoch zu beachten, dass die Rahmenbedingungen der Evaluation trotzdem kontrollierende Elemente haben, z. B. dass die Schulinspektion (außer in Schleswig-Holstein) nicht freiwillig ist.

In diesem Zusammenhang ist zu berücksichtigen, dass es auch Schulinspektionen gibt, die auf einen deskriptive Beschreibung der beobachteten Prozesse verzichten: *Nö, wir bewerten dis schon. Also wir- da bin ich auch- wie-. Die Debatte gibts immer wieder. Und es gibt auch Inspektorn, die Ihnen sagen würden: <Ich würd am liebsten nich bewerten>. So. Aber ich finde einfach, wir habn den Auftrag. Wir sind- Das is eben das, was denn, wenn man sich die Funktion anguckt, wird das sofort klar. Das is diese Spannung. Und ich finde, die muss man aushalten. So <lachen>. Das is- Man kann die nich lösen. Also man wird da nich zu na abschließenden Lösung komm. Im Sinne von: <So machn wirs jetz, und jetz is das richtig, das Verhältnis>. Das wird immer auch dazu führen, dass man das Gefühl hat: Mh, hier werd ich einer Sache nich gerecht, hier werd ich einer Schule an na bestimmten Stell nich gerecht. (.) Aber so is dis. Aber ich finde deswegen einfach zu sagen, wir bewerten nich mehr. Dann vereinfacht man das zu stark“ (Int. 13_43).*

8.2 Relevante Impulse beim Sensemaking

Im Folgenden werden verschiedene Ergebnisse des Sensemakings zu Inklusion bzw. zu den Bewertungsinstrumenten zu Inklusion thematisiert und dabei wird dargestellt, welche Rahmenbedingungen in Bezug auf diese Teilaspekte prägend wirkten. Es werden nun also die Ergebnisse aus dem ersten Teil der Analyse, den Deutungen von Inklusion (→ Kapitel 7), mit den bisherigen Ergebnissen zum Sensemaking (→ Abschnitt 8.1) miteinander verknüpft.

8.2.1 Sensemaking von Teilaspekten von Inklusion

Die **Entwicklung von Instrumenten in Bezug auf Inklusion** wurde in den einzelnen Bundesländern unter anderem durch Signale der Kultusministerien, schulgesetzliche Veränderungen sowie stattfindende Debatten in der Öffentlichkeit initiiert. Dabei wird bei der **inhaltlichen Gestaltung** primär auf die sonderpädagogische Expertise zurückgegriffen, die im eigenen Haus vorhanden ist sowie auf ihre Praxisbeobachtungen als Inspektor_innen. Die wissenschaftliche Literatur zu Inklusion hat sich bei der Entwicklung von Instrumenten als nicht hilfreich erwiesen. Darüber hinaus greifen die Instrumenteentwickler_innen auch bei der Entscheidung darüber, welche **Erwartungen an Schule** formuliert werden bzw. was als machbar angesehen wird, auf ihre Beobachtungen der Praxis bzw. eigener Praxistätigkeit zurück. Welche Signale sie von Seiten des Kultusministeriums und der schulischen Vertreter erhalten, kann ebenfalls von Relevanz sein. Es ist ebenfalls zu berücksichtigen, ob das Verfahren eher darauf abzielt die Einhaltung von Mindeststandards zu überprüfen (-Kontrollfunktion) oder Impulse für weitere Entwicklungsschritte zu formulieren (-Impulsfunktion). So können höhere Erwartungen formuliert werden, wenn die Impulsfunktion dominiert.

In diesem Zusammenhang spielt bei der **Entwicklung von Instrumenten** auch das jeweilige Verständnis von Inklusion eine Rolle. So ergibt sich aus Typ I „Anspruch auf Förderung“ (→ Kapitel 7.3), dass Mindeststandards formuliert werden, die Schulen mit Schüler_innen mit diagnostizierter Behinderung erfüllen müssen. Dagegen kann Inklusion im Fall von Typ III „Inklusion als Teilaspekt von guter Schule“ als Prozess verstanden werden, der schrittweise realisiert werden soll.

In Kapitel 7.2 ist aufgezeigt worden, dass sich die Bundesländer darin unterscheiden, wie sie **Instrumente zu Inklusion in das bestehende Instrumentenset integriert haben**: als Add-on oder als Querschnitt. Hierbei lassen sich Zusammenhänge zu dem Verständnis von Inklusion, das sich innerhalb der Schulinspektion durchgesetzt hat, herstellen: Bei einem engen Inklusionsbegriff liegt es nahe, Inklusion als Add-on zu etablieren, wenn es nur an denjenigen Schulen Anwendung finden soll, die Schüler_innen mit diagnostizierter Behinderung unterrichten. Dagegen legt ein weites Inklusionsverständnis nahe Instrumente zu Inklusion als Querschnitt zu integrieren.

Bei der Entscheidung darüber, ob die **Instrumente zu Inklusion verpflichtend oder freiwillig sind**, wird unter anderem auf die Rechtsprechung, die Relevanz des Themas sowie die Umsetzung von Inklusion im jeweiligen Bundesland verwiesen. So wird ausgehend von dem Argument, dass sich Deutschland rechtlich zur Umsetzung verpflichtet hat, argumentiert: *„Wenn man auf diese Behindertenrechtskonvention da zurückgreifen möchte – gedanklich. Dann muss dies eine Vorleistung der Schulen sein. Alle Schulen müssen vorbereitet sein und die Frage, ob hier ein behindertes Kind ist oder nicht, spielt hier auch überhaupt keine Rolle“* (Int. 17_17). Jedoch ergänzt diese Interviewperson, dass zugleich auch die Umsetzung von Inklusion im Bundesland sowie die Positionen der schulischen Vertreter berücksichtigt werden müssten, was eher dagegen spricht, dass Instrumente zu Inklusion verpflichtend sind. Dies ist auch konform mit den bereits dargestellten Positionen der schulischen Vertreter in Bezug auf die Höhe der Erwartungen, die an sie formuliert werden. Ein weiterer Umstand, der für die Entscheidung zwischen Wahl- und Pflichtbereich prägend wirkte, ist das Relevanzkriterium: *„Ich glaube unterschwellig ähm wollte man die wichtigen von den unwichtigen trennen. Und das hat aber nie jemand so formuliert. Denn das ist nicht vermittelbar, dass die anderen unwichtig sind“* (Int. 5_21). Dieses Argument wird nicht explizit im Zusammenhang mit den Instrumenten zu Inklusion genannt, jedoch

kann dieses Entscheidungskriterium auch darauf angewandt werden, ob die Instrumente zu Inklusion verpflichtend oder freiwillig sind.

8.2.2 Sensemaking zu Lernstandserhebungen

Ein weiterer relevanter Teilaspekt bei Inklusion ist die **Berücksichtigung von Schulleistungstests** in den Instrumenten der Schulinspektionen (für einen Überblick → Kapitel 7.1.1). Hierbei zeigte sich, dass in vielen Bundesländern nicht die Ergebnisse entscheidend sind, sondern der Umgang mit diesen. Wie bereits dargestellt, ist die Integration von Schulleistungstests nicht per se konträr zu dem hier festgelegten Verständnis von Inklusion, kann aber zu nicht-intendierten Wirkungen führen, die mit der Umsetzung von Inklusion nicht kompatibel sind. Als Argument für die stärkere Berücksichtigung von Lernstandserhebungen wurde im Rahmen eines informellen Gesprächs genannt, dass es sinnvoll wäre, wenn Schulinspektionen die Daten der Schulleistungstests nutzen, da sehr viele finanzielle Ressourcen in die Erhebung gesteckt wurden. Durch die Verwendung würden die Schulinspektionen die Vergleichsarbeiten als bildungspolitische Maßnahme stärken. Zu dieser Position konträr, argumentiert eine Person aus einem anderen Bundesland, dass die Berücksichtigung von Vergleichsarbeiten durch die Schulinspektion negative Folgen haben könnte: So wurde die Einführung von Vergleichsarbeiten im eigenen Bundesland von vielen sehr kritisch gesehen. Wenn die Schulinspektion die Ergebnisse berücksichtigen würde, könnte es ihrer Meinung nach verstärkt zu Gaming-Strategien von Seiten der Schulen führen, da sie versuchen könnten, die Ergebnisse zu manipulieren, z. B. indem sie den Schüler_innen mehr Zeit zur Verfügung stellen: *„Und da ist jetzt halt mühsam die Überzeugung [für die Vergleichsarbeiten im Land, ap] geschaffen worden und das würden wir komplett kaputt machen, wenn wir uns jetzt diese Daten nehmen und sagen, das ist gut oder nicht gut“* (Int. 1_47).

Darüber hinaus plädiert eine Interviewperson generell für eine stärkere Outputorientierung, die sich an Lernstandserhebungen im Querschnitt orientiert. *„Insofern würde mir da vorschweben, aber das ist meine individuelle Überzeugung, dass wir ein ganz, äh – ein ganz klares Outputmodell fahren würden [anonymisiert, ap]. Denkbar wäre das ne kleine Qualitätsagentur – die hier klare Qualitätserwartungen an Schule richtet und dann feststellt, inwiefern die erreicht werden und das zurückmeldet“* (Int. 8). Dahinter steht der Gedanke, dass die Schulen unter variierenden Rahmenbedingungen agieren und bestimmte Outputs mit unterschiedlichen Maßnahmen erreicht werden können. Die angestrebte Agentur könnte den einzelnen Schulen schulspezifische Unterstützung liefern, um sie dabei zu unterstützen die angestrebten Outputs zu erreichen (Int. 8). Hier wird eine starke Orientierung an evidenzbasierten Praktiken deutlich, die den Fokus auf Formen der Outputsteuerung legt und mit den aktuellen Verfahren der Schulinspektionen nicht kohärent sind.

Als weitere Argumente, welche gegen die Berücksichtigung der Vergleichsarbeiten sprechen, werden genannt: **Erstens, dass es die Setzung von oben gegeben hätte, diese nicht zu berücksichtigen** (Int. 6; 14_16). **Zweitens, ist kein kausaler Zusammenhang zwischen den Prozessen** – gemessen durch die Schulinspektionen – **und den Outputs** – gemessen durch die Lernstandserhebungen – herstellbar bzw. wird dieser auch nicht angestrebt. Gleichwohl schilderte eine Interviewperson den Fall, dass einzelne Schulen nachfragten, wie es denn sein könne, dass sie in den Vergleichsarbeiten gut abschneiden würden, aber auf der Prozessebene bei den Schulinspektionen schlecht (Int. 14_16). Dementsprechend formulieren die Schulen die Erwartung, dass sie, wenn sie auf der Outputebene, d.h. bei den Vergleichsarbeiten, gut abschneiden,

auch auf der Prozessebene, d.h. bei der Erhebung durch die Schulinspektionen, gute Ergebnisse aufweisen müssten. Dahinter steht der Gedanke, dass ein Ursache-Wirkungs-Zusammenhang zwischen den auf der Prozessebene gesammelten Daten von Seiten der Schulinspektion und den auf der Outputebene vorhandenen Daten durch die Vergleichsarbeiten besteht. Der Kommentar von schulischer Seite geht in eine ähnliche Richtung, wie die oben genannte Forderung nach stärkerer Outputorientierung. Sehr vereinfacht formuliert, steckt dahinter die polemisch formulierte Frage von Seiten der Schule: Was interessieren die Prozesse, wenn die Outputs stimmen? Dagegen argumentiert Biesta „Die Mittel, die wir in der Erziehung einsetzen, verhalten sich nicht neutral zu den Zielen, die wir erreichen wollen“ (vgl. Biesta 2011, S. 103). Demnach lässt sich auf den Bildungsbereich nicht die gleiche Evidenzlogik anwenden wie z. B. auf den medizinischen Bereich, wo eine Behandlung danach beurteilt wird, ob die Patient_innen danach wieder gesund sind.

Drittens spricht gegen die Berücksichtigung von Schulleistungstests, dass die Schulinspektionen im Fall von Deutschland (für England siehe Piezunka 2017) über keine Daten verfügen, um die Lernfortschritte von einzelnen Schüler_innen abbilden zu können. Während die vorherigen Begründungen, z. B. bildungspolitische Strategien, in keinem Zusammenhang zu Inklusion stehen, lässt sich in diesem Fall eine Verknüpfung herstellen: So wird hier die Notwendigkeit formuliert, dass Aussagen über Schulqualität nur möglich sind, wenn Daten zum Lernfortschritt im Längsschnitt vorhanden sind. Hier gibt es Anknüpfungspunkte zu dem in Kapitel 2 dargestellten Konzept „individuell bestmögliche Qualifizierung“. Jedoch stellen auch Daten im Längsschnitt keine Operationalisierung des hier verwendeten Verständnisses von Inklusion dar, weil dieses sich von der Erwartung einer normalen Lernkurve distanziert. Diese ist jedoch notwendig, um Längsschnittdaten einordnen zu können.

Allgemein lässt sich in Hinblick auf die Berücksichtigung von Inklusion durch das Beispiel Schulleistungstests Folgendes herausarbeiten: Inklusion als Konzept sowie Schulleistungstests repräsentieren zwei unterschiedliche Verständnisse von guter Schule, die in der konkreten Praxis nur begrenzt miteinander kompatibel sind bzw. konträre Signale setzen können (nähere Ausführung → Kapitel 2). In der Begründung für bzw. gegen Schulleistungstests taucht bei der Entwicklung der Instrumente jedoch kein wertebasiertes Argument auf, das auf ein bestimmtes Qualitätsverständnis verweist. Stattdessen sind es strategische Gründe, die im Mittelpunkt stehen. Dies legt nahe, dass die Instrumente der Schulinspektion häufig mehrere Verständnisse von guter Schule repräsentieren und das Ergebnis von Kompromissen darstellen. In Bezug auf die Entwicklung von Instrumenten macht dieses Beispiel deutlich, dass es häufig pragmatische Gründe sind, die zu der Entscheidung für oder gegen bestimmte Setzungen führen. Dabei wird – nicht nur bei diesem Beispiel – selten die normative Komponente des Prozesses explizit angesprochen.

8.3 Sensemaking von einzelnen Fällen

In den Interviews zeigen sich verschiedene Deutungen von Inklusion (nähere Ausführung → Kapitel 7.3). Im Folgenden wird das Sensemaking auf Einzelfallebene beschrieben, um das Interagieren der Rahmenbedingungen abbilden zu können. Die vorgestellten Fälle sind exemplarisch. So können es auch andere Rahmenbedingungen sein, die zu ähnlichen Deutungen führen (nähere Ausführung → Kapitel 8.3.3).

8.3.1 Frau Elsner als Vertreterin von Typ I

Typ I „Inklusion als Anspruch auf Förderung“: Danach wird Inklusion als gemeinsamer Unterricht von Schüler_innen mit und ohne diagnostizierter Behinderung verstanden. Bei den Schulinspektionen besteht Handlungsbedarf, da diese sicherstellen sollen, ob Schüler_innen mit diagnostizierter Behinderung die Unterstützung erhalten, die sie auch an Förderschulen erhalten haben.

Frau Elsner³ arbeitete jahrelang als Sonderpädagogin. Neben ihrer Tätigkeit als Lehrkraft war sie zusätzlich als Fortbildnerin zuständig und hat sich in diesem Zusammenhang auch mit schulrechtlichen Fragestellungen auseinandergesetzt. In ihrer Tätigkeit bei der Inspektion fühlt sie sich primär für das Thema Inklusion verantwortlich bzw. wird ihr dieser Aufgabenbereich von ihren Kolleg_innen zugeschrieben.

Für die Entwicklung von Instrumenten hat sie sich zu Beginn sehr intensiv mit dem „Index für Inklusion“ auseinandergesetzt. Daraus entstand die Idee eigene Indikatoren zu entwickeln, die angelehnt an den Index die Dimensionen „Haltung“; „organisatorischer Rahmen“ sowie „unterrichtliche Praxis“ abbilden (zum Vergleich: „Kulturen“, „Strukturen“ und „Praktiken“ im „Index für Inklusion“). Dabei stand die Situation einzelner Schülergruppen zunächst nicht im Vordergrund. Stattdessen ging es darum, ob Schulen realisiert haben, dass die Umsetzung von Inklusion ein Umdenken erfordert, dass sich in vielen Bereichen des schulischen Lebens widerspiegelt, z. B. individuelle Differenzierung im Unterricht. Bei der konkreten Formulierung von Instrumenten konnte Frau Elsner neben dem „Index für Inklusion“ ihre Expertise als Sonderpädagogin einbringen. So macht sie im Rahmen des Interviews am Beispiel der Differenzierung deutlich, dass sie als Sonderpädagogin – im Vergleich zu Regelschullehrkräften – über mehr Erfahrung in diesem Bereich verfügt und daher auch höhere Anforderungen gegenüber Schulen formuliert.

Während des Entwicklungsprozesses der Instrumente kam sie von ihrem ersten Entwurf ab: *„Da kriegt diese Inklusion so ein Gewicht und eigentlich wäre Inklusion ja am besten, man merkt nicht“*. So hätte ihr erster Entwurf den Schulen deutlich signalisiert, in welchem Ausmaß ein Umdenken bezüglich inklusiver Haltungen erforderlich ist. Dies hätte dazu führen können, dass die Schulen sich abwehrend gegenüber den Instrumenten zu Inklusion positionieren. Gleichwohl thematisiert sie im Laufe des Interviews immer wieder, dass sie die Schulen als Akteure wahrnimmt, die zum großen Teil noch nicht verstanden hätten, welche Implikationen mit der Umsetzung von Inklusion einhergehen: *„Wie sich das entwickelt und ich glaube, was generell auch noch so ne Schwierigkeit ist, dass wir da Erwartungen haben und die Schulen haben teilweise - glaube ich - noch garnicht richtig erkannt, dass dies letztendlich ihre Aufgabe ist“*. Demnach ist es ihr Anliegen Schulen in Bewegung zu bringen, aber Abwehrreaktionen zu vermeiden, indem sie deren fehlende Erfahrung zu Inklusion bei der Formulierung von Instrumenten berücksichtigt: *„Ich glaube, weil das schon bei uns war – gemessen auch in anderen Bundesländern – [Nennung des Bundeslandes] so ein bisschen hinterher hinkte“*. „Aber irgendwann wurde dann aber klar. Das ist irgendwie zu groß“. Eine weitere Herausforderung war, dass eine Orientierung an bundesländerspezifischen Rechtsvorgaben bzw. eine rechtliche Legitimation nicht möglich war, weil die Anpassungen des Schulgesetzes zeitgleich zu den Veränderungen der Bewertungsinstrumente stattfanden.

Darüber hinaus war die Frage von Relevanz, wie die Instrumente zu Inklusion in das Instrumententsetz integriert werden sollten. So sah sie zu Beginn der Auseinandersetzung die Instrumente

3 Hierbei wird auf Quellenangaben verzichtet, um die Anonymität der Gesprächspartner zu gewährleisten.

zu Inklusion als Add-on an und innerhalb ihrer Schulinspektion stand zur Debatte, ob den Schulen freigestellt wird, ob sie in diesem Bereich evaluiert werden möchten. Jedoch kamen sie und ihre Kolleg_innen im Laufe der Zeit von diesem Vorgehen ab: So zeichnete sich bildungspolitisch ab, dass sich die Schulen der Umsetzung von Inklusion nicht mehr entziehen konnten: *„Ich glaube dann wäre es komisch, wenn wir das als extra Merkmal hätten abgefragt“*. Demnach distanzierte sie sich von der Idee die Instrumente zu Inklusion als einen (optionalen) Zusatzbereich hinzuzufügen.

Die geringe Erfahrung der Schulen, bildungspolitische Entwicklungen im Bundesland sowie die fehlende rechtliche Legitimation führten daher zu einer neuen Herangehensweise bei der Entwicklung von Instrumenten zu Inklusion: Erstens wurde die Anzahl der inklusionsorientierten Instrumenten reduziert. Diese stellten innerhalb des Instrumentensets keinen eigenen Bereich mehr dar und es wurde für diejenigen Schulen, die Schüler_innen mit diagnostizierter Behinderung beschulen, verpflichtend. Dabei lag der neue Fokus darauf, dass Schüler_innen mit Behinderung an der Regelschule die gleiche Unterstützung erhalten sollen, die sie auf der Förderschule erhalten haben: *„Inklusion ohne Frage. Aber großes Fragezeichen für - glaube ich - alle Sonderpädagogen: <Kriegen die Schüler in der Inklusion das, was sie brauchen?>“*. Frau Elsner bezeichnet das neue Vorgehen als sonderpädagogischen Weg, da es nun primär um die Situation von Schüler_innen mit Behinderung und deren Anspruch auf besondere Unterstützung geht. So konnte sie sich mit diesem Vorgehen als Sonderpädagogin insofern identifizieren, als sie durch ihre berufliche Ausrichtung einen besonderen Bezug zu dieser Schülergruppe hat.

Trotzdem weist dieses enge Verständnis von Inklusion von ihrem Verständnis ab: Sie macht deutlich, dass sie unter Inklusion eigentlich den Umgang mit heterogenen Lerngruppen versteht. Auch wenn in den Instrumenten der Inklusionsbegriff anders verwendet wird, hat sie den Eindruck, dass ihr Verständnis von Inklusion als Erwartung im gesamten Instrumentenset bereits unter dem Konzept „Umgang mit Heterogenität“ berücksichtigt ist. Darüber hinaus kann sie sich mit dem Fokus auf bestimmte Mindeststandards gut identifizieren: *„Also ich bin jetzt - obwohl ich vorher auf diesem ganz ganz Großen und irgendwie auch - ja - das schon auch toll gefunden hätte - bin ich Gott froh, dass wir da ganz weggekommen sind“*. Durch das neue Vorgehen ist nun sichergestellt, dass die Schüler_innen, die sie früher unterrichtet hat, weiterhin die Unterstützung erhalten, die sie benötigen. Ihre Position ist nun auch kohärent mit den Signalen, die die Schulinspektion aus dem Kultusministerium erhalten hat. Dort wird ein enger Inklusionsbegriff vertreten.

Zusammenfassend wirkt es so, als wenn Frau Elsner zu Beginn des Sensemakings zwischen Typ III „Inklusion als Beschreibung von guter Schule“ und Typ I „Inklusion als Anspruch auf Förderung“ verortet werden kann (→ Abschnitt 7.3). Als die Entwicklung von Instrumenten konkreter wird, bewegt sie sich stärker zu einem Verständnis, das durch Typ I repräsentiert wird. Dies hat mehrere Gründe: Erstens hat sie den Eindruck, dass die Berücksichtigung von weiteren Differenzlinien in den Instrumenten mit dem Konzept „Umgang mit Heterogenität“ abgedeckt ist. Demnach taucht der weite Inklusionsbegriff für sie in den Instrumenten ihres Bundeslands auf, aber wird nicht als „Inklusion“ bezeichnet. Zweitens erhofft sie sich, dass durch den Fokus auf Schüler_innen mit diagnostizierter Behinderung und auf eine niedrigere Anzahl von Instrumenten von Seiten der Schulen weniger Widerstand entsteht. Wenn die Instrumente für die Schulen machbar erscheinen, könnte dies dazu führen, dass sie in Hinblick auf die Umsetzung von Inklusion aktiv werden.

Bei Frau Elsner werden auf der Ebene „persönliche Attribute“ ihre Erfahrung als Sonderpädagogin sowie ein Verständnis von guter Schule, wonach im schulischen Setting niemand verloren

gehen sollte, sichtbar. Des Weiteren versteht sie sich selbst als jemand, der dabei mitwirken möchte, dass die Umsetzung von Inklusion im eigenen Bundesland weiter vorangetrieben wird. Zwischen ihr und ihren Kolleg_innen fanden darüber hinaus Aushandlungsprozesse statt, die von den Signalen durch das Kultusministerium und schulischen Vertretern geprägt waren. Der Fokus auf Schüler_innen mit diagnostizierter Behinderung in den finalen Instrumenten stellt für sie einen Kompromiss dar, der die Situation im jeweiligen Bundesland repräsentiert. Dieser kam zustande, als im Laufe der Zeit die Signale aus dem Kultusministerium, die antizipierten

Tab. 8.2: Sensemaking bei Frau Elsner

Dimensionen des Sensemakings	Konkrete Rahmenbedingungen	Ausprägungen bei Frau Elsner
Persönliche Attribute des Sensemakers	Qualifikation und Praxiserfahrung	Sonderpädagogische Expertise sowie Lehrerfahrung im gemeinsamen Unterricht sowie als Fortbildnerin Zuständigkeit für das Thema
	Verständnis von Schulqualität	Umgang mit Differenzierung ist ein Teilaspekt von guter Schule
Situativer Kontext	Forschung	Die Dimensionen „Kulturen, Strukturen und Praktiken“ des Index‘ für Inklusion.
	Austausch mit Kolleg_innen	Machbarkeit der Instrumente bzw. antizipieren Positionen von schulischen Vertretern
	Umsetzung von Inklusion im Bundesland	Bundesland verfügt bislang über wenig Erfahrung zu Inklusion
	Funktion	Kontrollfunktion legt Formulierung von Mindeststandards nahe.
	Schulische Vertreter	Überforderung der Schulen soll vermieden werden
	Kultusministerium	Verwendung eines engen Inklusionsbegriffs, keine Signale in Bezug auf konkrete Anforderung Alle Schulen müssen sich mit der Umsetzung von Inklusion auseinandersetzen.
Rechtliche Vorgabe	BRK	Initiierung von Prozessen
	Bundesländer-spezifische Gesetzesvorgaben	Waren zum Zeitpunkt der Entwicklung der Instrumente noch nicht vorhanden.

Reaktionen von Schulen sowie die Erfahrungen der Schulen mit der Umsetzung von Inklusion stärker berücksichtigt wurden.

8.3.2 Herr Pehla als Vertreter von Typ II

Typ II „Inklusion als Herausforderung für Schulen“ (nähere Ausführung → Abschnitt 7.3): Wie bei Typ I wird auch hier Inklusion als gemeinsamer Unterricht von Schüler_innen mit und ohne diagnostizierte Behinderung verstanden. Jedoch wird Inklusion als eine Herausforderung für Schulen wahrgenommen. Das Interesse der Schulinspektion ist es herauszufinden, ob die Schulen in Bezug auf die Umsetzung von Inklusion aktiv werden, aber zugleich wollen sie nicht zu hohe Anforderungen an diese formulieren.

Herr Pehla kann Typ II zugeordnet werden. Er war ca. 20 Jahre lang als Lehrer tätig und danach Referent im Kultusministerium eines Bundeslandes, das er selbst als *„wertkonservativ und trotzdem innovativ“* beschreibt und sich durch ein gegliedertes Schulsystem kennzeichnet. Als Anhänger eines selektiven Schulsystems argumentiert er: *„Und dieses Bildungswesen war in allen Bildungsvergleichen top, sehr erfolgreich – gerade auch, was die Gymnasien anbelangt, aber nicht nur“*. Des Weiteren ist er in Bezug auf Inklusion unsicher, inwiefern die Umsetzung überhaupt machbar ist: *„Wir stehen in so einem Spannungsverhältnis, weil, also ich war ja selber lange Lehrer, 20 Jahre, und meine eben da auch genügend Praktiker zu sein und meine persönliche Meinung ist, dass was Hattie erwartet, was auch jetzt erwartet wird im ganzen Kontext von Differenzierung oder gar Individualisierung, Inklusion ist für die meisten Lehrer im Schulalltag nur schwer erreichbar, nur unbefriedigend erreichbar.“* Demnach geht er auch davon aus, dass die schulischen Vertreter sich gegen Inklusion aussprechen werden: *„ich vermute in gleichem Maße [...] werden ähm werden [...], anonymisiert] unsere Schulen [...], anonymisiert] auf die Barrikaden gehen und sagen das wollen wir nicht. Wir wollen eine exklusive Schule und keine inklusive. Aber wir werden sehen“*.

Als Befürworter eines selektiven Bildungssystems und aufgrund der Frage der Machbarkeit positioniert sich Herr Pehla daher negativ gegenüber der Umsetzung von Inklusion: *„Ich selber bin auch sehr skeptisch, um nicht zu sagen kritisch der Inklusion gegenüber eingestellt. Also gegenüber persönlich. Dienstlich muss ich ja eine andere Position einnehmen.“* Dies erklärt möglicherweise auch seine fehlenden Kenntnisse. So stellt er fest: *„Also ich kenn mich bei Inklusion überhaupt nicht aus...“*. Dies hat auch zur Folge, dass er bei der konkreten Entwicklung der Instrumente zu Inklusion nicht beteiligt war.

Dennoch diagnostiziert er aus dem Kontext der Schulinspektion heraus Handlungsbedarf in Bezug auf die Entwicklung von Instrumenten zu Inklusion: Zum einen aufgrund der rechtlichen Situation: *„Ich meine, dass eben die Bundesregierung sich verpflichtet hat, vielleicht diese Konvention eben umzusetzen. Und ich glaube, also dadurch ist das sozusagen die Rechtsnorm geworden in Deutschland, die wir erfüllen müssen“*. Des Weiteren gibt es auch von Seiten des Ministeriums eine deutliche Positionierung zu Inklusion: *„Ähm und dann gibts aber natürlich neben dieser Formalien, gibts ja die politische Ebene und die oberste Spitze [anonymisiert, ap], die vielleicht auch von seinem weltanschaulichem und bildungspolitischem Hintergrund her, sowas wie Inklusion voranbringen will“*. Demnach sieht er sich verpflichtet, als Mitarbeiter der Schulinspektion sich mit dem Thema Inklusion auseinanderzusetzen.

Bei der Integration der entwickelten Instrumente zu Inklusion in das bestehende Instrumentenset findet er sich mit seiner Positionierung insofern wieder, als die Instrumente zu Inklusion im Instrumentenset eine geringe Rolle spielen: So stellen sie ein Add-On zum bestehenden Instrumentenset dar. Herr Pehla hat generell zwar die Position, dass die Schulinspektion die

Tab. 8.3: Sensemaking bei Herrn Pehla

Dimensionen des Sensemakings	Konkrete Rahmenbedingungen	Ausprägungen bei Herrn Pehla
Persönliche Attribute des Sensemakers	Qualifikation und Praxiserfahrung	Erfahrung als Lehrkraft in einem konservativen Bundesland und Referent im Kultusministerium
	Verständnis von Schulqualität	Anhänger eines selektiven Bildungssystems
Situativer Kontext	Forschung	Hattie usw. formulieren zu hohe Anforderungen
	Austausch mit Kolleg_innen	Zulieferung von Instrumenten von Expert_innen
	Schulische Vertreter	Kritische Positionierung gegenüber Inklusion
	Kultusministerium	Befürwortung der Umsetzung von Inklusion
Rechtliche Vorgabe	BRK	Aus der BRK ergibt sich Handlungsbedarf in Bezug auf die Entwicklung von Instrumenten zu Inklusion.

Aufgabe hat „*einem Kern zu definieren und für mehr einheitliche Standards zu sorgen*“, aber die Durchsetzung von Standards zu Inklusion ist ihm persönlich kein Anliegen.

Zusammenfassend zeigt sich, dass sich bei Herrn Pehla Bezüge zu seinem Verständnis von guter Schule herstellen lassen. Als Anhänger eines segregierten Bildungssystems, das er als Praktiker kennengelernt hat, steht er der Umsetzung von Inklusion eher kritisch gegenüber: Er bezweifelt, dass Inklusion zu einer Verbesserung des Bildungssystems bezüglich Leistungsentwicklung führen würde. Dennoch setzt er sich mit Inklusion auseinander: So ergibt sich durch die rechtlichen Bedingungen und die Signale, die vom Ministerium ausgehen, die Notwendigkeit Instrumente zu Inklusion in das bestehende Instrumentenset zu integrieren. Jedoch sieht er es als eine Maßnahme an, die die Schulen nun umsetzen müssen und die deswegen auch von den Schulinspektionen berücksichtigt werden sollte. Er ist überzeugt davon, dass die Umsetzung von Inklusion auch von den schulischen Vertretern nicht gewollt ist.

8.3.3 Herr Plate und Frau Steigleder als Vertreter von Typ III

Typ III „Inklusion als Teilaspekt von guter Schule“ (nähere Ausführung → Abschnitt 7.3): Danach wird Inklusion als gelungener Umgang mit heterogenen Gruppen verstanden. Inklusion wird von Seiten der Schulinspektion berücksichtigt, weil es eine Beschreibung von guter Schule darstellt.

Im Folgenden wird das Sensemaking von Herrn Plate und Frau Steigleder in Bezug auf Inklusion dargestellt. Dabei zeigt sich, dass sie ein ähnliches Verständnis von Inklusion haben. Gleichwohl unterscheiden sie sich darin, wie die Umsetzung von Inklusion gemessen werden sollte.

Entwicklung von Bewertungsinstrumenten zu Inklusion: Herr Plate: Er ist Grundschullehrkraft mit sonderpädagogischer Zusatzqualifikation. Was sein Inklusionsverständnis angeht, verweist er auf seine Tätigkeit in der Wissenschaft im Bereich der Integrations- bzw. Inklusionsforschung als prägende Zeit und fühlt sich der Bremer Schule rund um Georg Feuser zugehörig. Er sieht sich selbst als jemanden, der sich bereits seit Jahren um die Umsetzung von Inklusion bzw. Integration bemüht: *„Und wir haben so in den 90er Jahren schon sehr viel für Inklusion getan und diesen Begriff eigentlich nicht genutzt“*. Nach seiner Zeit in der Forschung hat er in seiner Tätigkeit als Lehrkraft weiterhin versucht die theoretischen Erkenntnisse in der Praxis umsetzen.

Seine Expertise zu Integration bzw. Inklusion war auch der Grund, warum er zur Schulinspektion gekommen ist: *„Und das war der Punkt, wo dann gesagt wurde – wir brauchen dann nen paar neue Leute, die dann – sage ich mal – diesen anderen Stil von Unterricht, der Bandbreite abdeckt, auch mal selber unterrichtet haben. Also wissen, wovon sie sprechen“*. Dabei ging es bei seiner Einstellung nicht primär um Inklusion, sondern um die Themenbereiche Individualisierung und Förderschule. *„Ich sag mal. Ich habe dann aus meiner Intention natürlich dann, weil ich entsprechend geprägt wurde, über die Bremer Schule dann das Thema Inklusion hier im Haus forciert.(...) Ja und. Ich bin jetzt sozusagen der Inklusionsexperte hier“*. (...) *„Weil die natürlich gemerkt haben – die stoßen mit ihrem Hilbert Meyer an Grenzen. Inklusion ist jetzt aber [...] Keiner wusste so richtig, was das eigentlich war“*.

Bei der Entwicklung von Instrumenten greift er insbesondere auf den „Index für Inklusion“ zurück. In diesem Zusammenhang ist zwischen zwei Instrumentenversionen zu unterscheiden: Zum einen gibt es die Instrumente, die Herr Plate in Anlehnung an sein Inklusionsverständnis entwickelt hat, aber die aktuell keine Anwendung finden: *„Es gibt ein Wahlthema, das habe ich entwickelt. In der Schublade. Was die Schulen es wählen könnten, wenn wir es anbieten. Also wie gesagt, es liegt noch in der Schublade. Das hebt auf den Index für Inklusion ab“*. Zum anderen gibt es einzelne Indikatoren zu Inklusion im Instrumentenset, welches aktuell in seinem Bundesland Anwendung findet. Jedoch sind diese Instrumente nicht explizit als „inklusionsorientiert“ etikettiert.

Die Tatsache, dass sein favorisierter Entwurf aktuell keine Anwendung findet, begründet er unter anderem damit, dass die Schulen in ihrer Entwicklung noch nicht so weit seien. Ausschlaggebend sind jedoch Positionen seiner Kolleg_innen gegenüber Inklusion. Indem er auf ihre Vorerfahrung als Praktiker verweist, begründet er ihre kritische Position gegenüber Inklusion: *„Es waren doch schon viele kommunikativ. Aber doch einige auch, die sagen. Der Hilbert Meyer ist so eigentlich das Wahre und mein Paradigma – und daran wird jetzt nicht gerüttelt. Und was ihr hier über Inklusion erzählt und gemeinsamer Lerngegenstand und inklusive Lernarrangements – das ist alles halligalli und das wird ja nie funktionieren können“*. Einen Verweis auf die Kolleg_innen gibt es auch, als es darum geht, ob Instrumente zu Inklusion ein Wahlthema sein sollten. Dafür spricht, dass es seinen Kolleg_innen an Expertise fehlt: *„Ich habe das ja vorhin beschrieben. Ehemalige Schulräte zum großen Teil, Sekundarschullehrer, die auch ihre eigene Unterrichtspraxis erst vorheriges Jahr hatten – also mit diesen Dingen überhaupt nicht in Berührung gekommen sind. Insofern doch viele Kollegen Berührungsängste haben und hatten. Ich habe das über mehrere Jahre schon abbauen können – zum Teil. Aber es es gibt da schon noch Vorbehalte, die da auch beißen. Also nicht nur Hilbert Meyer ist gut, sondern wir können das nicht“*. Demnach fehlt es den Kolleg_innen an Expertise, die Umsetzung von Inklusion an Schulen evaluieren zu können: *„Wir müssen das ein Stück weit in der Schublade lassen, weil – wenn wir das anbieten – wir gehen an die Schulen. Wir können das nicht leisten. Wir sind nicht kompetent genug“*.

In Bezug auf die Funktionen der Schulinspektion argumentiert er: *„weil unser Verständnis unseres Verfahrens ist nicht so sehr Bewertung in diesem klassischen Sinne – das macht ihr gut, das macht ihr schlecht, sondern eher im Sinne von Evaluation als gemeinsamer Verständigungsprozess als ein darüber Nachdenken, was tun wir, wo stehen wir und daraus entwickeln, was sind unsere Schritte“*. Demnach steht für ihn die Impulsfunktion von Schulinspektionen im Mittelpunkt und er distanziert sich während des Interviews immer wieder von einer möglichen Kontrollfunktion von Schulinspektionen. Aus dieser Perspektive heraus, würde es in Bezug auf die Messinstrumente naheliegen, dass Schulen frei wählen können, ob sie zu Inklusion evaluiert werden möchten. Hierbei wird jedoch deutlich, dass ihm die Umsetzung von Inklusion auf der persönlichen Ebene ein sehr wichtiges Anliegen ist und er dieses Thema gerne sowohl bei den Schulen als auch innerhalb der Schulinspektion stärker verankern und es deshalb zu einem verpflichtenden Thema machen möchte: *„ja, also, da muss man wirklich nur darüber nachdenken, ob das der richtige Weg ist ein Wahlthema anzubieten und es an anderen Stellen der Struktur [anonymisiert, ap] nicht zu benutzen oder ob da eine komplett neue Architektur ranmüsste, die dann insgesamt heißt inklusive Schulentwicklung und wir untersetzen das jetzt mit unterschiedlichen Themenbereichen, also das muss bedacht werden“*.

In Bezug auf die Entwicklung von Instrumenten ist sich Herr Plate darüber hinaus noch unsicher, wie er sich gegenüber dem Kultusministerium positionieren soll. Generell nimmt er dieses nicht als besonders beeinflussend wahr: *„Die sind vom [Kultusministerium, ap] – sage ich mal – die sind relativ zurückhaltend. Die schreiben uns da nicht vor also <wir wollen jetzt, dass ihr das und das in den Fokus nehmt>“*. Gleichwohl verweist er auf die Position des Kultusministeriums, um zu begründen, warum Inklusion als Begriff in den Instrumenten kaum auftaucht: *„Wenn ich jetzt einfach mal schaue, wie wird Inklusion hier im Land kommuniziert. Was ist das? Dann muss ich sagen. Dann muss ich sagen: Leider sind wir inzwischen an dem Punkt in [Nennung des Bundeslands], wo Inklusion gleich gesetzt wird mit <gemeinsamem Unterricht> und Betreuung von Kindern mit sonderpädagogischem Förderbedarf. [...] Insofern unter diesem Hintergrund dieser sehr engen Sicht auf Inklusion und ich sage: Das ist es nicht. Das dürfte nicht als Inklusion bezeichnet werden. Habe ich mich also hier in unserem Fachbereich sehr stark dafür gemacht, dass vor dem Hintergrund dieser Denke der Begriff Inklusion nicht verwendet wird“*. In Bezug auf zukünftige Entwicklungen sieht er die Instrumente der Schulinspektion als eine Möglichkeit an, ein Inklusionsverständnis zu etablieren, welches von dem des Kultusministeriums abweicht. Demnach denkt er darüber nach den Inklusionsbegriff stärker in den Instrumenten zu verankern und hierbei auf sein Verständnis von Inklusion zurückzugreifen und nicht auf das Verständnis des Kultusministeriums: *„Im Sinne von <wir müssen auch unseren Beitrag leisten ein anderes Inklusionsverständnis zu entwickeln>“*.

Zusammenfassend zeigt sich bei Herrn Plate, dass insbesondere seine Tätigkeit im Bereich der Integrations- bzw. Inklusionsforschung prägend für die Herausbildung seines Verständnisses von Inklusion war. Bei der konkreten Entwicklung von Instrumenten orientiert er sich am „Index für Inklusion“, jedoch kann er sich mit seinem Entwurf an Instrumenten aktuell nicht durchsetzen. Dies begründet er primär damit, dass das Verständnis von Schulqualität seiner Kolleg_innen von seinem abweicht und es darüber hinaus bei diesen an Expertise zu Inklusion fehlt. Im Gegensatz zu anderen Instrumenteentwickler_innen ist bei Herrn Plate auffällig, dass schulische Vertreter, die Umsetzung von Inklusion im eigenen Land sowie die Position des Kultusministeriums bei der Entwicklung von Instrumenten eine geringe Rolle spielen. Hierbei ist jedoch zu beachten, dass sein Entwurf an Instrumenten aktuell keine Anwendung findet. In Bezug auf die tatsächlich verwendeten Instrumente im Instrumentenset waren sowohl das

Tab. 8.4: Sensemaking bei Herrn Plate

Dimensionen des Sensemakings	Konkrete Rahmenbedingungen	Ausprägungen bei Herrn Plate
Persönliche Attribute des Sensemakers	Qualifikation und Praxiserfahrung	Tätigkeit als Grundschullehrkraft mit sonderpädagogischer Zusatzqualifikation; Tätigkeit im Bereich der Integrations- und Inklusionsforschung
	Verständnis von Schulqualität	Demokratische Schule; Differenzierung im Unterricht
Situativer Kontext	Forschung	Orientierung am „Index für Inklusion“
	Kolleg_innen innerhalb der Schulinspektion	Fehlende fachliche Expertise zu Inklusion; Kritische Position zu Inklusion
	Funktionen von Schulinspektionen	Sieht sich nicht als rechtsprüfende Instanz; Schulinspektion soll den Schulen Impulse zur Schulentwicklung geben
	Kultusministerium	Keine inhaltliche Positionierung zu den konkreten Inhalten der Instrumente; Vertreten engen Inklusionsbegriff
	Schulische Vertreter	Antizipiert viel Handlungsbedarf in Bezug auf die Umsetzung von Inklusion
Rechtliche Vorgabe	BRK	Keine Erwähnung

Inklusionsverständnis des Kultusministeriums als auch die Positionen seiner Kolleg_innen prägend. Während des Interviews wird ebenfalls deutlich, dass rechtliche Vorgaben zu Inklusion kaum relevant waren. Dies wird damit begründet, dass die Schulinspektion nicht für die Überprüfung der Einhaltung von Rechten zuständig sei.

Frau Steigleder: Bei Typ III bietet es sich an neben Herrn Plate noch das Sensemaking von Frau Steigleder zu beschreiben: Sie hat vor ihrer Tätigkeit in der Schulinspektion in verschiedenen Bundesländern als Lehrerin an Förderschulen gearbeitet. Sie arbeitete auch als Leiterin einer Förderschule, die Teil einer kooperativen Schule war. Im Anschluss kam sie an die Schulinspektion, wo sie aufgrund ihrer sonderpädagogischen Expertise für den Themenbereich Inklusion mit zuständig war bzw. ist.

Ihr eigenes Verständnis von guter Schule begründet sie mit ihren Beobachtungen aus der Praxis als Inspektorin: *„Also das liegt halt einfach auch daran, weil man glaub ich, wenn man Schulinspektion betreibt und einfach viele Schulen sieht, dann sieht man viel (.) Realität. Ne? (.) Und- (.) und- und sieht eigentlich, dass- (.) dass nur das ständige Bestreben nach ähm (.) ähm einem positiven Umgang mit Vielfalt n Lösungsweg sein kann“*. Demnach ist für sie der gelungene

Tab. 8.5: Sensemaking bei Frau Steigleder

Dimensionen des Sensemakings	Konkrete Rahmenbedingungen	Ausprägungen bei Frau Steigleder
Persönliche Attribute des Sensemakers	Qualifikation und Praxiserfahrung	Sonderpädagogische Expertise; Beobachtungen aus der Praxis als Inspektorin
	Verständnis von Schulqualität	Umgang mit Vielfalt als Gewinn
Situativer Kontext	Kolleg_innen innerhalb der Schulinspektion	Inklusion sei der richtige Weg; Intensive Auseinandersetzung mit Inklusion
	Forschung	Aargauer Bewertungsraster und Index für Inklusion
	Schulische Vertreter	Spannungsfeld zwischen Ernstnehmen und Fordern
Rechtliche Vorgabe	BRK	Keine Erwähnung

Umgang mit Heterogenität maßgeblich im Hinblick auf Schulqualität. Daran anknüpfend versteht sie unter Inklusion: „*Inklusive Schule wäre für mich jede gute Schule. Wenn – wenn es gelingt mit Vielfalt umzugehen, dann äh- (.) dann gelingt Inklusion*“.

Bei der Entwicklung von Instrumenten erwähnt sie verschiedene Rahmenbedingungen: Zum einen wurde immer wieder der Austausch mit anderen Abteilungen aus der Bildungsverwaltung bzw. des Ministeriums gesucht: „*Dass wir da immer gut mitkriegen (.) ähm, was treibt die Leute da grade um, passt das was wir tun dazu. [...] und da sehn wir zu, dass wir da gut auch informiert sind. Und äh (.) mitgesch- also möglichst mitgestalten*“. Zum anderen hat ein intensiver Austausch mit ihren Kolleg_innen stattgefunden. Hier waren zunächst unterschiedliche Positionen vertreten. Jedoch bedeutet dies nicht, dass bei ihren Kolleg_innen kritische Haltungen gegenüber der Umsetzung von Inklusion vorhanden waren: „*Ich glaub einfach schon, dadurch, dass wir alle sehr viel sehr unterschiedliche Schuln sehn, is hier schon- schon eigentlich- eigentlich äh Konsens, (.) dass Inklusion der Weg sein muss*“. Gemeinsam mit ihren Kolleg_innen hat sie dann auch die Bewertungsinstrumente aus dem ersten Zyklus mit den Indikatoren aus dem „Aargauer Bewertungsraster“ und dem „Index für Inklusion“ abgeglichen: „*bilden Kriterien das ab, was Inklusion ausmacht, habn wir für uns entschieden: Ja, dis is so*“. Dies hat zur Folge, dass zwischen dem ersten und zweiten Zyklus keine zusätzlichen Instrumente in Bezug auf Inklusion entwickelt wurden, weil Inklusion – nach ihrem Verständnis der gelungene Umgang mit Heterogenität – bereits durch die Instrumente abgedeckt wird. In diesem Fall hat demnach eine Umdeutung des vorhandenen Instrumentensets stattgefunden und keine Neuentwicklung.

Für Frau Steigleder spielen schulische Vertreter eine relevante Rolle, jedoch, was die Formulierung von Erwartungen betrifft, ist sie der Meinung, dass Schulen häufig auf Ausreden zurückgreifen, wenn es um die Machbarkeit geht: „*Ähm, (.) ich glaube Schuln könn das erfülln*“.

Ich glaube, wenn- wenn Schuln meinen, ähm, (.) diese Bewertungskriterien seien unerfüllbar, dann-dann- (.) <seufzen> dann sind das auch kleine Fluchten“. Auf der anderen Seite gesteht sie jedoch ein, dass die Schulinspektion als Akteur für diese Haltung auch bildungspolitisch bestraft worden ist [keine weitere Ausführung aus Datenschutzgründen]. Des Weiteren argumentiert sie im Zusammenhang mit konkreten Beispielen, z. B. Differenzierung im Unterricht oder Förderplangestaltung, dass sie dies nicht fordern können, weil sie sonst folgende Rückmeldung von Seiten der Schulen erhalten würden: *„Und dann ham wir gesagt, ja so- so- sone Rückmeldung sollten wir ma na Schule geben. Die würden uns vom Hofjagen“.* Demnach argumentiert sie, dass es in solchen kritischen Situationen, in denen Schulen die Erwartungen als zu hoch empfinden, es notwendig wäre eine rechtliche Legitimation zu haben, die bei vielen Teilaspekten von Inklusion in ihrem Bundesland nicht existiert.

Zusammenfassend zeigt sich, dass Frau Steigleder Inklusion als gelungenen Umgang mit Heterogenität definiert und dies ihrem Verständnis von guter Schule entspricht. Dabei verweist sie immer wieder auf ihre Beobachtungen aus der Praxis, um dieses Verständnis von guter Schule zu begründen. Innerhalb ihres Kollegiums gibt es zwar Diskussionen zum Verständnis von Inklusion, aber die Notwendigkeit des Konzepts als solches wird nicht in Frage gestellt. Auch wenn Inklusion innerhalb der Schulinspektion Zustimmung erfährt, führt dies nicht dazu, dass neue Instrumente zu Inklusion in das bestehende Instrumentenset integriert werden, sondern vorhandene Instrumente werden umgedeutet und als „inklusionsorientiert“ interpretiert.

Vergleich zwischen Herrn Plate und Frau Steigleder: Im Vergleich zwischen Herrn Plate und Frau Steigleder wird deutlich, dass beide unter Inklusion einen gelungenen Umgang mit Heterogenität verstehen und dies ihrem Verständnis von guter Schule entspricht. Dies ist auch die Begründung dafür, warum Inklusion in den Instrumenten der Schulinspektion berücksichtigt werden sollte (nähere Ausführung zu Typ III → Kapitel 7.3). Bei der Analyse der Interviews wird jedoch sichtbar, dass sie auf unterschiedliche Erfahrungen zurückgreifen, wenn es darum geht, wie sie zu diesem Verständnis gekommen sind: Herr Plate verweist auf seine Tätigkeiten im Wissenschaftsbereich, die ihn stark geprägt haben. Dagegen hat Frau Steigleder als Lehrkraft und als Inspektorin den Eindruck gewonnen, dass ein gelungener Umgang mit Heterogenität maßgeblich dafür ist, dass Schulen – nach ihrer Einschätzung – funktionieren. Auch wenn beide das gleiche Verständnis von Inklusion aufweisen, führt dies zu unterschiedlichen Implikationen. Ausgehend von seinem Inklusionsverständnis entwickelt Herr Plate auf der Grundlage des „Index für Inklusion“ neue Instrumente, die aber unter anderem aufgrund seiner Kolleg_innen in der Schulinspektion keine Anwendung finden. Dagegen gleichen Frau Steigleder und ihre Kolleg_innen das vorhandene Instrumentenset mit dem „Index für Inklusion“ ab und kommen zu dem Ergebnis, dass das Instrumentenset bereits ihr Verständnis von Inklusion abdecken würde. Nun könnte auf den ersten Blick argumentiert werden, dass beide Personen bei der Entwicklung von Instrumenten ähnlich vorgehen, aber dass im Fall von Frau Steigleder das bestehende Instrumentenset bereits die Instrumente, die beiden nach dem „Index für Inklusion“ vorschweben, enthält. Dies ist jedoch nicht der Fall. So entsprechen die Instrumente, die im Bundesland von Frau Steigleder Anwendung finden, nicht den Instrumenten, die Herrn Plate vorschweben (keine nähere Ausführung aus Datenschutzgründen).

Zusammenfassend kann festgestellt werden, dass bei Herrn Plate und Frau Steigleder unterschiedliche Rahmenbedingungen wirkten, aber diese trotzdem zu einem ähnlichen Verständnis von Inklusion gekommen sind. Aufgrund weiterer Rahmenbedingungen aus dem situativen

Kontext hat dies jedoch nicht dazu geführt, dass ähnliche Instrumente zu Inklusion entwickelt wurden bzw. diese in den jeweiligen Bundesländern Anwendung fanden.

8.4 Kollektive Aushandlungsprozesse

Im Fokus der Arbeit steht das individuelle Sensemaking bei der Entwicklung von Instrumenten. Diese können jedoch nicht getrennt von kollektiven Aushandlungsprozessen innerhalb der Schulinspektion betrachtet werden. So prägen Positionen und Stimmungen innerhalb des Teams von Instrumenteentwickler_innen das individuelle Sensemaking bzw. das Sensemaking der einzelnen Instrumenteentwickler_innen fließt ein, wenn im Rahmen von kollektiven Aushandlungsprozessen Instrumente entwickelt und formuliert werden.

Im Rahmen der Arbeit gab es bereits erste Andeutungen darüber, inwiefern kollektive Aushandlungsprozesse bei der Entwicklung von Instrumenten eine Rolle spielen, z. B. konnte sich Herr Plate mit seinem Verständnis von guter Schule innerhalb seiner Schulinspektion nicht durchsetzen, weil seine Kolleg_innen sich dazu mehrheitlich kritisch positionierten: *„Schwierig, mh - sagen wir nicht schwierig, sondern langwierig. Ja. Ich sag mal nen Paradigmenwechsel herbeizuführen. Das ist relativ schwierig. Das ist bei Schulen so. Das ist bei solchen Einheiten so. (...) Paradigmenwechsel geht nur mit ner biologischen Lösung, also es gibt Leute, die vertreten ein Paradigma und werden damit groß - reden das mit weiter“* (Herr Plate). In seinem Fall hätte er hohes Interesse daran, dass Inklusion anders und stärker in den Instrumenten verankert wird. Im oben beschriebenen Fall von Herrn Pehla ist es dagegen genau umgekehrt: Inklusion wird von ihm eher kritisch gesehen und es ist der Druck von Seiten des Kultusministeriums sowie die Positionen seiner Kolleg_innen, die dazu führen, dass Inklusion in den Instrumenten berücksichtigt wird und somit seine Herangehensweise, die wahrscheinlich nahelegen würde, dass Inklusion nicht thematisiert wird, sich in den Instrumenten nicht wiederfindet. In diesem Fall ist das Ergebnis von kollektiven Aushandlungsprozessen, dass neue Instrumente zu Inklusion in das bestehende Instrumentenset integriert werden und dies von Herrn Pehla akzeptiert werden muss. Des Weiteren zeigt sich in vielen Fällen (z. B. Int. 15_10; 3_12), dass innerhalb der Schulinspektion ähnliche Positionierungen zu Inklusion vorhanden sind: *„Es gibt keine grundsätzlich unterschiedliche- keine grundsätzlich divergierende Haltung dazu, glaub ich“* (15_10). Dies wird in einem Fall damit erklärt, dass bei der Personalauswahl innerhalb der Schulinspektion Personen ausgesucht werden, die ähnliche Vorstellungen haben.

Darüber hinaus geht es im Rahmen von kollektiven Aushandlungsprozessen häufig um die Frage, ob ein enger oder weiter Inklusionsbegriff verwendet werden soll und inwiefern Inklusion – im Falle eines weiten Inklusionsbegriffs – von dem Konzept „Umgang mit Heterogenität“ abgegrenzt werden kann (Int. 17_39; 1_16). Wie bereits oben dargestellt spielen dabei noch weitere Rahmenbedingungen wie Vorgaben vom Kultusministerium und individuelle Überzeugungen eine Rolle.

Es lässt sich festhalten, dass das individuelle Sensemaking zu Inklusion bzw. zu den jeweiligen Instrumenten erste Hinweise darauf gibt, wie diese entstanden sind. Als Ergänzung ist es jedoch notwendig die kollektiven Aushandlungsprozesse innerhalb des Teams der Instrumenteentwickler_innen herauszuarbeiten.

Die Notwendigkeit Kompromisse einzugehen oder Adaptionen vorzunehmen, wird anhand der folgenden Beispiele deutlich. So spielen bei der Entwicklung von Instrumenten häufig bundeslandspezifische Rahmenbedingungen eine Rolle, die sich in den jeweiligen Qualitätsrahmen widerspiegeln: *„Ich denk mal - so ein Qualitätsrahmen ist in erster Linie ein normativer*

Bezugsrahmen, d.h. es ist aushandelbar, was da drinsteht im Grunde genommen. Auch wenn man sich an wissenschaftlichen - äh - Fragestellungen orientiert oder Ergebnissen orientiert, äh, kann man darüber diskutieren, was man unter Qualität verstehen will“ (Int. 17_6; siehe auch 10_19) beziehungsweise argumentiert eine andere Befragte, dass man die Vorgaben der Wissenschaft durchaus nochmal auf die Machbarkeit überprüfen müsste (Int. 1_43). Des Weiteren argumentieren andere Interviewpersonen, dass die Frage nach Schulqualität auch davon abhängig wäre, welche Akteursgruppe befragt wird (Int. 11_17; 8; 16_23) beziehungsweise, dass es von bildungspolitischen und historischen Rahmenbedingungen abhängt (Int. 2_8). So zeigt sich der bildungspolitische Einfluss beispielsweise bei der Inklusion: *„Der einzige Bezugspunkt war im Grunde genommen die Ergebnisse wissenschaftlicher Untersuchungen, die belegen, dass bestimmte Qualitätskriterien Einfluss auf die Schülerleistungen nehmen, aber das ist natürlich nicht so, dass es dann rein aufgeht. Da gibt es natürliche politische Interessen, die ja auch - die auch irgendwie ihr Recht haben. Das war insbesondere bei der Inklusion“* (Int. 17_4; siehe auch 4_4). Darüber hinaus wird das Suchen nach Kompromissen auch darin deutlich, dass eine Interviewperson beschreibt, dass unterschiedliche Anforderungen an Schulen formuliert werden, die sich gegenseitig widersprechen *„Äh es ist ja sogar auch ein ähm, eine ... eigentlich ein innerer Widerspruch in dem Auftrag, der an die Schule [anonymisiert, ap] gestellt ist. Einerseits soll sie inklusiv arbeiten, andererseits sind sie exklusiv und sie sollen eigentlich auch exklusiv sein“* (Int. 8_33). Die oben dargestellten Zitate machen demnach deutlich, dass die Entwicklung von Instrumenten häufig ein Aushandlungsprozess ist, der mit einer Vielzahl von Kompromissen endet und abhängig davon ist, welche Rahmenbedingungen im jeweiligen Fall für relevant erachtet werden. Des Weiteren zeigen sie erneut, dass im tatsächlichen Handeln die Idee eines allgemein anerkannten Verständnisses von Schulqualität nicht aufrecht erhalten werden kann.

8.5 Einzelfallübergreifende Handlungsmuster

Im Rahmen dieses Kapitels wurde dargestellt, welche Rahmenbedingungen das Sensemaking zu Inklusion prägten, wie die Rahmenbedingungen die Deutung von Teilaspekten von Inklusion beeinflusst haben und wie diese auf Einzelfallebene miteinander interagieren. Hierbei ließen sich verschiedene Handlungsmuster herausarbeiten, die bei der Entwicklung von Bewertungsinstrumenten zu Inklusion von Relevanz waren und über den Individualfall hinausgehend gelten.

Erstens zeigte sich als Muster die starke Relevanz von eigenen Praxiserfahrungen als Lehrkraft sowie Beobachtungen aus der Praxis als Inspektor_in. Diese waren maßgeblich für die eigene Positionierung zu Inklusion sowie für die Frage danach, inwiefern die Umsetzung von Inklusion als machbar angesehen wird. In Bezug auf die Entwicklung von Instrumenten spiegelte sich dies – in Interaktion mit anderen Rahmenbedingungen – darin wider, welche Relevanz dem Thema Inklusion beigemessen wurde und welche Erwartungen gegenüber Schulen formuliert wurden. Darüber hinaus spielten diese Rahmenbedingungen insbesondere bei der inhaltlichen Formulierung von Instrumenten und der Entscheidung darüber, welche Teilaspekte von Inklusion Eingang in die Instrumente finden sollen, eine Rolle.

Zweitens zeigte sich interviewübergreifend als Muster das Streben nach Akzeptanz. In diesem Zusammenhang sind verschiedene Akteure zu nennen, die von den einzelnen Interviewpartnern für wichtig erachtet wurden: Schulische Vertreter, das Kultusministerium sowie die Schulaufsicht. Dagegen spielten beispielsweise die Eltern, Schüler_innen sowie Lobbyverbände bei der Entwicklung von Instrumenten kaum eine Rolle. In den Interviews variierte es, welchen

Akteuren besondere Bedeutung zugeschrieben wurde sowie welche Impulse von Seiten der jeweiligen Akteure in Bezug auf Inklusion ausgingen. Es zeigte sich jedoch einzelfallübergreifend, dass die Interessen dieser Akteure von Relevanz waren. Es gab nur Unterschiede darin, in welchem Ausmaß und wie auf diese Interessen eingegangen wurde. In Hinblick auf die Entwicklung von Instrumenten zu Inklusion spiegelt sich dieses Handlungsmuster unter anderem darin wider, welche Erwartungen gegenüber Schulen im Hinblick auf die Umsetzung von Inklusion formuliert werden und wie der Inklusionsbegriff verwendet wird.

Als drittes Handlungsmuster zeigt sich, dass rechtliche Vorgaben wie die BRK die Auseinandersetzung über Instrumente zu Inklusion initiiert haben. Jedoch haben sich aus den rechtlichen Vorgaben fast nie inhaltlichen Hinweise im Zusammenhang mit den Instrumenten ergeben. Demnach wurde die Auseinandersetzung mit Inklusion häufig damit begründet, dass sich Deutschland durch die BRK zur Umsetzung eines inklusiven Bildungssystems verpflichtet hätte. Bei der konkreten Entwicklung von Instrumenten wurde jedoch nicht auf konkrete Inhalte der rechtlichen Vorgaben zurückgegriffen. Dies wird unter anderem damit begründet, dass sich die Schulinspektion nicht in der Funktion sieht, die Einhaltung von Rechtsvorgaben zu überprüfen. Dies ist aber nicht kohärent mit der Argumentation, dass manche Instrumenteentwickler_innen aufgrund von rechtlichen Vorgaben begonnen haben, Inklusion als Konzept in den Instrumenten zu verankern. Zusammenfassend zeigt sich, dass diese Handlungsmuster über den Einzelfall hinausgehend sichtbar werden.

9 Diskussion und Ausblick

Im Rahmen der Arbeit wurde näher betrachtet, wie Inklusion in den Instrumenten der Schulinspektionen berücksichtigt wurde. In einer zweistufigen Analyse wurde dabei zunächst untersucht, ob relevante Konzepte von Inklusion, die sich aus dem in der Arbeit verwendeten Inklusionsverständnis¹ (nähere Ausführung → Kapitel 2) ergeben, in den Messinstrumenten aufgeführt und wie diese operationalisiert werden. Des Weiteren wurde herausgearbeitet, wie Inklusion von den einzelnen Schulinspektionen gedeutet wird. Hierfür wurden zum einen die Instrumente analysiert, die von Seiten der Schulinspektionen als „inklusionsorientierte“ Instrumente gekennzeichnet wurden, und zum anderen wurden die Instrumenteentwickler_innen nach ihren Inklusionsverständnissen befragt. Zusätzlich wurde im Rahmen der Arbeit die Entwicklung der Instrumente zu Inklusion in den Blick genommen. So konnte durch die empirische Analyse der Interviews gezeigt werden, welche Rahmenbedingungen die Deutung von Inklusion sowie die Entwicklung der Instrumente prägen und welche Handlungsmuster bei der Entwicklung der Instrumente sichtbar werden. Bei der Analyse der Messinstrumente konnten die Instrumente der Bundesländer miteinander verglichen werden. Dies war aus Datenschutzgründen bei der Analyse der Interviews nicht möglich.

Das folgende Diskussionskapitel greift die durchgeführten Analysen in vier Abschnitten auf:

1. Berücksichtigung von Inklusion in den Instrumenten der Schulinspektionen (→ Abschnitt 9.1.1);
2. Operationalisierung von Inklusion in den Instrumenten der Schulinspektionen (→ Abschnitt 9.1.2);
3. Deutungen der Instrumenteentwickler_innen zu Inklusion (→ Abschnitt 9.1.4);
4. Relevante Rahmenbedingungen bei der Entwicklung der Messinstrumente (→ Abschnitt 9.2).

Zu Beginn jedes Teilabschnitts werden die Ergebnisse der empirischen Analyse aus dem siebten und achten Kapitel kurz zusammengefasst. Anschließend werden diese diskutiert und hierbei werden Bezüge zu bestehender Forschung hergestellt und die Implikationen der Ergebnisse besprochen.

9.1 Berücksichtigung von Inklusion

Zunächst wurden die Messinstrumente der Schulinspektionen mit dem im Rahmen dieser Arbeit festgelegten Verständnis von Inklusion verglichen. Diese Auswertung stellt eine Ergänzung zu der bislang einzig vorhandenen Studie dar, die sich mit einer ähnlichen Fragestellung auseinandersetzt: So wird bei Degenhardt (2008) der Fokus auf die Frage gelegt, inwiefern Begriffe wie „Integration/ Inklusion“ explizit in den Instrumenten der Schulinspektionen genannt werden. Durch das von Degenhardt gewählte Vorgehen kann jedoch nicht abgebildet werden, ob Konzepte von Inklusion möglicherweise in den Instrumenten operationalisiert wurden, aber der Begriff als solcher nicht explizit genannt wurde.

¹ „Individuell bestmögliche Qualifizierung“, „politische Partizipation aller“ und „Erfahren von intersubjektiver Anerkennung“ als Ziele von Inklusion.

9.1.1 Konzepte in Bezug auf Inklusion

Kurzzusammenfassung der Ergebnisse aus der Analyse: In Bezug auf das Ziel von Inklusion „individuell bestmögliche Qualifizierung“ zeigt sich, dass in fast allen Bundesländern Formen der inneren Differenzierung durch diverse Messinstrumente der Schulinspektionen abgefragt werden. Hierbei wird jedoch selten explizit von Differenzierung auf Individualebene gesprochen. Stattdessen zielen die Indikatoren häufig auf Differenzierung zwischen verschiedenen Gruppen innerhalb einer Lerngruppe ab, z. B. auf unterschiedliche Aufgabenformate für leistungsschwache und -starke Schüler_innen. Als weiteres Ziel von Inklusion wurde im Rahmen der Arbeit die politische Partizipation aller Schüler_innen festgelegt. Hierzu hat sich in der empirischen Analyse gezeigt, dass dieser Aspekt in den Instrumenten aller Bundesländer operationalisiert wird. Jedoch unterscheiden sich die Bundesländer darin, wie politische Partizipation gemessen wird. Darüber hinaus zeigte sich in der empirischen Analyse zu politischer Partizipation, dass nur in wenigen Bundesländern (HH; BE; BW; SH) überprüft wird, inwiefern alle Schüler_innen die Möglichkeit haben, den schulischen Alltag mitzugestalten. Auch in den oben aufgeführten Bundesländern wird politische Partizipation nur in Bezug auf einzelne Teilgruppen, z. B. in Bezug auf Schüler_innen mit Behinderung, überprüft. Bei der Analyse des Inklusionsziels „Erfahren von intersubjektiver Anerkennung“ zeigen sich ähnliche Muster wie bei dem eben betrachteten Ziel „politische Partizipation“: Der Aspekt taucht in Form von wertschätzendem Umgang in den Messinstrumenten aller Bundesländern auf, jedoch unterscheiden sie sich darin, wie sie den Aspekt messen und welche Relevanz sie dem Thema beimessen. Des Weiteren zeigt sich auch bei dem Ziel „Erfahren von intersubjektiver Anerkennung“, dass bei der Operationalisierung die Situation von einzelnen Schüler_innen kaum in den Blick genommen wird.

Neben den Zielen von Inklusion wurde im Rahmen der Arbeit untersucht, welche Differenzlinien bzw. Heterogenitätsdimensionen in den Instrumenten explizit genannt werden und wie diese operationalisiert werden. Dies ist insofern von Relevanz, weil das in der Arbeit verwendete Verständnis von Inklusion einen besonderen Fokus auf die Differenzlinien legt, die in Bezug auf Teilhabe und Diskriminierung von Bedeutung sind. Dabei wurde im Rahmen einer deduktiven Analyse untersucht, inwiefern die Differenzlinien genannt werden, bei denen es schon empirische Belege dafür gibt, dass diese in Bezug auf Bildungsungleichheiten bedeutsam sind, z. B. sozioökonomischer Status. Zugleich erlaubt das Verfahren der qualitativen Inhaltsanalyse nach Gläser und Laudel (2009), dass durch eine induktive Vorgehensweise weitere Differenzlinien in den Blick genommen werden können, z. B. „private Ausnahmesituationen“.

Im Rahmen der Analyse zeigte sich, dass „Leistung“ in den Instrumenten der Schulinspektionen im Vergleich zu anderen Differenzlinien eine sehr dominante Rolle einnimmt. Dagegen wird die Heterogenitätsdimension „Behinderung“ in vielen, aber nicht in allen Bundesländern explizit in den Instrumenten aufgeführt. Neben „Behinderung“ und „Leistung“ wurde untersucht, ob weitere Differenzlinien wie „Gender“, „sozioökonomischer Status“, „private Ausnahmesituationen“ sowie „Fremdwahrnehmung als Nicht-Deutsche“ als Differenzlinien aufgeführt werden. Diese Heterogenitätsdimensionen tauchen jedoch nur in einigen wenigen Bundesländern auf, z. B. in Hamburg und Berlin. Darüber hinaus gibt es noch andere Heterogenitätsdimensionen, z. B. sexuelle Orientierung oder Körpergewicht, die in keinem der Bundesländer auftauchen, obwohl es empirische Belege dafür gibt, dass diese Differenzlinien in Bezug auf Teilhabe und Diskriminierung von Relevanz sind (z. B. Helbig und Jähnen 2013; Hartmann 2012).

Es lässt sich abschließend festhalten, dass eine Vielzahl von Differenzlinien in den Messinstrumenten der Schulinspektionen nicht explizit benannt wird. Daran anknüpfend werden auch

der Aspekt der Intersektionalität sowie die Unterscheidung zwischen Fremd- und Selbstzuschreibung in keinem Bundesland von Seiten der Schulinspektionen explizit angesprochen. Generell zeigt sich, dass sich die Schulinspektionen der Bundesländer darin unterscheiden, bei welchen gesellschaftlichen Funktionen von Schulen eine inklusive Perspektive eingenommen wird, welche Differenzlinien hierbei ausdrücklich genannt und wie diese operationalisiert werden.

Berücksichtigung des Inklusionsverständnisses dieser Arbeit in den Instrumenten: Bei der Betrachtung der Ergebnisse stellt sich die Frage, inwiefern in den Instrumenten der Schulinspektion jenes Verständnis von Inklusion operationalisiert wurde, welches im Rahmen dieser Arbeit als normative Vergleichsfolie dient. So gibt es eine Vielzahl von Messinstrumenten, die überprüft, inwiefern bestimmte Aspekte für einzelne Teilgruppen gewährleistet sind. Jedoch gibt es kein Bundesland, in dem in Bezug auf alle untersuchten Ziele das Individuum in den Blick genommen wird. Demnach bleiben alle Bundesländer unter den Anforderungen, die sich aus dem in dieser Arbeit verwendeten Inklusionsverständnis ergeben. Hierbei ist zu berücksichtigen, dass bei den Schulinspektionen Erhebungsformen, z. B. der Fragebogen für Schüler_innen und Eltern, existieren, die es ermöglichen würden, die individuelle Perspektive zu erfassen. So zeigt sich im Fall von Hamburg und Bayern, dass der Aspekt „Anerkennung als Individuum“ erfasst wird: Jedoch sind es die Lehrkräfte, die im Rahmen des Lehrerfragebogens befragt werden, inwiefern sie den Eindruck haben, in ihrer Individualität gesehen und anerkannt zu werden. Dagegen wird im Schülerfragebogen von beiden Bundesländern nicht thematisiert, inwiefern auch Schüler_innen den Eindruck haben, dass sie in ihrer Individualität gesehen und wahrgenommen werden. Nur im Land Berlin wird dieser Aspekt auch in Bezug auf Schüler_innen thematisiert. Auch wenn die individuelle Perspektive in den Instrumenten selten eingenommen wird, lässt sich festhalten, dass in vielen Bundesländern ein erstes Fundament gelegt wurde, da häufig die Situation von einzelnen Schülergruppen in den Blick genommen wird.

Bezüglich des Inklusionsziels „individuell bestmögliche Qualifizierung“ wurde bereits im Rahmen der theoretischen Diskussion deutlich, dass es in der bildungswissenschaftlichen Literatur wenige Überlegungen dazu gibt, wie dieser Aspekt durch Messungen auf der Outputebene berücksichtigt werden könnte. Die Herausforderung, jenen Aspekt zu messen, wird auch in der Studie von Sowada und Dederling von einem Inspektor thematisiert:

„Wir haben das dann vermutet, aber wenn man ganz ehrlich ist, kann man Lernzuwachs als Beobachter, der zwanzig Minuten in einer Stunde ist, nicht beobachten, weil wir die Ausgangssituation nicht wissen. Und nur, weil Kind plötzlich schlau guckt, können wir nicht davon ausgehen, dass es jetzt einen Lernzuwachs hatte, ne?“ (Sowada und Dederling 2016, S. 191).

Auch wenn die Inspektionen auf andere Daten zurückgreifen würden, z. B. Längsschnitterhebungen wie es bei OFSTED der Fall ist, ist es nicht möglich zu evaluieren, ob dies der bestmögliche Fortschritt für das jeweilige Individuum ist (vgl. Bradbury und Roberts-Holmes 2017). So kann durch die Daten zwar abgebildet werden, ob Lernfortschritte vorhanden sind. Die Bewertung des Lernfortschritts kann jedoch nur durch einen Vergleich erzielt werden, z. B. mit der restlichen Lerngruppe oder Schüler_innen mit ähnlichen Ausgangslagen. Es ist jedoch mithilfe der Daten nicht möglich, Aussagen darüber zu treffen, ob dies der bestmögliche Fortschritt für die jeweilige Person ist. Demnach liegt es aus Inklusionsperspektive nahe darauf zu verzichten, Daten auf der Outputebene zu erheben und stattdessen zu erfassen, inwiefern die Gelingensbedingungen für individuelle Förderung auf der Prozessebene gegeben sind. Ein Fokus

auf die Prozessebene ist auch kohärent mit den Messinstrumenten aller Schulinspektionen in Deutschland.

Nachdem im Hinblick auf die Ziele von Inklusion deutlich wurde, dass die individuelle Perspektive selten in den Blick genommen wird, wird im nächsten Schritt ein weiterer Aspekt des hier verwendeten Inklusionsverständnisses näher betrachtet: die explizite Nennung von Differenzlinien. Hierbei ist zu berücksichtigen, dass in der Praxis sowie in der Forschung häufig unterschiedliche Bezeichnungen verwendet werden, um Gruppen zu beschreiben, die z. T. mit ähnlichen Herausforderungen konfrontiert sind, z. B. „Menschen mit Migrationshintergrund“ oder „Menschen mit Deutsch als Zweitsprache“ oder „Menschen, die von anderen als Nicht-Deutsche“ wahrgenommen werden. Um in den Blick nehmen zu können, welche Bezeichnungen von Seiten der Schulinspektionen verwendet werden, wurde im Rahmen der Arbeit mit Sammelbegriffen gearbeitet. Demnach wurden alle genannten Differenzlinien einem der abstrakten Sammelbegriffe zugeordnet. Dies ermöglicht eine Strukturierung des Materials und daran anschließend eine Auseinandersetzung mit der Frage, welche Gruppen von den einzelnen Schulinspektionen in den Blick genommen werden und welche Implikationen mit den gewählten Bezeichnungen einhergehen. Beispielsweise tauchen unter dem Sammelbegriff „Behinderung“² bei den Messinstrumenten der Schulinspektionen unter anderem die Bezeichnungen „sonderpädagogischer Förderbedarf“, „Anspruch auf sonderpädagogisches Förderangebot“ oder „körperliche Beeinträchtigung“ auf (nähere Ausführung in Bezug auf Behinderung folgt in diesem Abschnitt).

Im Hinblick auf die Differenzlinie „Leistung“ lässt sich festhalten, dass deren dominante Rolle aufgrund der Qualifizierungs- und Selektionsfunktion von Schule (vgl. Fend 2006) wenig überraschend ist. Bei der Operationalisierung von Leistung wird in den Instrumenten häufig auf dichotome Unterscheidungen zurückgegriffen, indem beispielsweise innerhalb einer Lerngruppe zwischen leistungsschwachen und -starken Schüler_innen unterschieden wird. Dies ist in Anlehnung an das hier verwendete Verständnis von Inklusion kritisch zu sehen: Durch diese Form der Operationalisierung entsteht der Eindruck, dass die jeweilige „Leistungsfähigkeit“ zeitlich und fächerübergreifend stabil sei und Schüler_innen entweder leistungsstark oder -schwach seien (vgl. Prengel 2014). Dadurch wird nicht sichtbar, dass alle Zuschreibungen sozial konstruiert sind und es zeitlich und kontextabhängig variieren kann, wie die Fähigkeit von einzelnen Schüler_innen einzuschätzen ist. Des Weiteren wird dadurch auch nicht berücksichtigt, dass bei der Leistungseinschätzung durch Lehrkräfte häufig auch andere Aspekte eine Rolle spielen, die unabhängig sind von den Fähigkeiten, die gemessen werden sollen, z. B. Zuschreibungen von Leistungsfähigkeit aufgrund des Etiketts „Migrationshintergrund“ (z. B. Sprietsma 2009).

Bei der Differenzlinie „Behinderung“ ist Folgendes festzuhalten: Die Nennung als solche ist insofern nicht überraschend, als im öffentlichen Diskurs unter Inklusion häufig der gemeinsame Unterricht von Schüler_innen mit und ohne Behinderung verstanden wird. Jedoch unterscheiden sich die Bundesländer darin, welche Bezeichnungen sie für „Behinderung“ verwenden. Ausgehend von der Frage, inwiefern die BRK in die Instrumente der Schulinspektion implementiert wurde, lässt sich festhalten, dass die von den Inspektionen verwendeten Bezeichnungen häufig nur eine Teilgruppe derjenigen Schüler_innen abbilden, deren Rechte durch die BRK geschützt werden (vgl. Piezunka et al. 2018). So wird in keinem Bundesland die Gruppe von

² Im Rahmen dieser Arbeit dient „Behinderung“ als Sammelbegriff für verschiedene Umschreibungen, die von Seiten der Schulinspektion genutzt werden, um die Gruppe von Schüler_innen zu beschreiben, die aufgrund von langfristigen körperlichen oder psychischen Beeinträchtigungen an der gleichberechtigten gesellschaftlichen Teilhabe behindert werden (vgl. Bielefeldt 2009).

Schüler_innen in den Messinstrumenten explizit aufgeführt, die nach SGB VIII oder XII eine Behinderung aufweist und daher Anspruch auf besondere Unterstützung hat (in Bezug auf „besondere Unterstützung“ siehe Piezunka et al. 2016). Darüber hinaus zeigt sich, dass lediglich in der niedersächsischen Schulinspektion die Gruppe der Schüler_innen mit Behinderung als heterogen wahrgenommen wird. Thematisiert wird dort, dass sich in Abhängigkeit von der diagnostizierten Beeinträchtigung unterschiedliche Anforderungen für den schulischen Kontext ergeben. Im Gegensatz dazu werden in den anderen Ländern Schüler_innen mit Behinderung – wenn sie überhaupt als Gruppe explizit aufgeführt werden – als homogene Gruppe operationalisiert. Demnach bestätigt diese Arbeit die Ergebnisse der Studie von Degenhardt: Er hat untersucht, inwieweit die Gruppe von Schüler_innen mit Sehbehinderung in den Instrumenten der Schulinspektionen explizit genannt wird und hierbei festgestellt, dass diese Gruppe nur in einem Fall (Rheinland-Pfalz) und im Zusammenhang mit Förderschulen genannt wurde (vgl. Degenhardt 2008, S. 15).

In Hinblick auf das im Rahmen dieser Arbeit verwendete Inklusionsverständnis, ist darauf hinzuweisen, dass es über die Anforderungen hinausgeht, die in Art. 24 BRK formuliert werden: Es wurden nicht nur die Schüler_innen, die unter den Schutz der BRK fallen, in den Blick genommen, sondern die gesamte Schülerschaft. So beziehen sich die Anforderungen, die sich aus Art. 24 BRK ergeben, nur auf eine Teilgruppe der Schülerschaft, nämlich Menschen mit diagnostizierter Behinderung. Mit Blick auf die Implementation von Art. 24 BRK lässt sich Folgendes festhalten: Durch die empirische Analyse wurde deutlich, dass die Situation von Menschen mit diagnostizierter Behinderung bei den betrachteten Funktionen von Schule, z. B. Qualifizierung und Partizipation, nicht immer explizit aufgeführt wird bzw. in manchen Bundesländern diese Gruppe überhaupt nicht genannt wird. Demnach lassen Daten, die mithilfe der Messinstrumente der Schulinspektionen erhoben werden, nur begrenzt Aussagen darüber zu, inwiefern im jeweiligen Bundesland auf Einzelschulebene Art. 24 BRK umgesetzt wurde. Hierfür wären weitere Messinstrumente notwendig.

Implikationen der konkreten Operationalisierung der Teilaspekte: Im Rahmen dieser Arbeit geht es unter anderem darum, wie Teilaspekte von Inklusion in den Messinstrumenten von Schulinspektionen operationalisiert werden. In diesem Zusammenhang stellt sich die Frage, welche Implikationen mit der jeweiligen Operationalisierung einhergehen.

So lässt sich beispielhaft am Ziel „Erfahren von intersubjektiver Anerkennung“ abbilden, welche Wirkung quantitative Indikatoren haben: So wird der wertschätzende Umgang zwischen verschiedenen schulischen Akteuren durch verschiedene Indikatoren im Unterrichtsbeobachtungsbogen erhoben. Dabei wird meistens auf eine zwei- oder vierstufige Antwortskala zurückgegriffen. Beispielsweise im bayerischen Unterrichtsbeobachtungsbogen: „Die Lehrkraft geht freundlich mit den Schülern um“ (UBB_BY). Als Antwortskala ist „(eher) ja“ und „(eher) nein“ möglich.³ Um mögliche Implikationen der Operationalisierung sichtbar zu machen, bietet es sich an, auf die INTAKT-Studie (für einen Überblick siehe Fallarchiv Universität Kassel 2018) zu verweisen: Im Rahmen der Studie wurden die pädagogischen Beziehungen zwischen Fachkräften und jungen Erwachsenen analysiert und in diesem Zusammenhang untersucht, wie häufig es zu seelischen Verletzungen im pädagogischen Setting kommt. Im Rahmen der Studie wurde sichtbar, dass es im pädagogischen Setting immer wieder zu ambivalenten Situationen kommt, welche „Mehrdeutigkeiten und Widersprüche aufweisen“ (Tellisch 2016). Es ist beispielsweise eine ambivalente Situation, wenn Bemerkungen von Lehrkräften positiv intendiert

3 Darüber hinaus gibt es noch „keine Angabe möglich“ sowie ein Anmerkungsfeld.

sind, aber die Personen, die die Situation von außen beobachten, sich unsicher sind, ob die Bemerkungen von Seiten der Schüler_innen als verletzend oder als ermutigend wahrgenommen werden. Des Weiteren kann die Einschätzung darüber, was als freundlicher Umgang verstanden wird, zwischen den Inspektor_innen variieren.

Darüber hinaus wird bei dem oben aufgeführten Indikator aus dem bayerischen UBB die Interaktion zwischen Lehrkraft und der ganzen Gruppe in den Blick genommen. Es kann jedoch davon ausgegangen werden, dass die Qualität der Beziehungen zwischen Lehrkraft und einzelnen Schüler_innen variiert (vgl. Prengel und Winklhofer 2014). Dieser Aspekt wird durch das gewählte Item und die dazugehörige Antwortskala nicht sichtbar. Quantitative Indikatoren lassen außerdem auch keine Aussagen darüber zu, was mögliche Ursachen bzw. Gelingensbedingungen für die Qualität pädagogischer Beziehungen sind. Demnach zeigen sich in Bezug auf diesen quantitativ erhobenen Indikator Phänomene, die Espeland und Sauder (2007) mit dem Konzept der Kommensurierung beschreiben: So kann durch numerische Werte die Komplexität von sozialen Phänomenen nur begrenzt abgebildet werden, weil durch die Form der Datenerhebung Kontextwissen verloren geht (nähere Ausführung → Kapitel 4).

Darüber hinaus zeigte sich bei den Instrumenten zu **politischer Partizipation**, aber auch bei den anderen Zielen von Inklusion, dass diese bundesländerübergreifend nicht einheitlich operationalisiert werden. Dies ist insofern nicht überraschend: So fehlt es beispielsweise an Standards, die definieren, was unter politischer Partizipation in Bildungssettings zu verstehen ist (vgl. Deutsche Gesellschaft für Demokratiepädagogik e.V. 2014). Daher kann von Seiten der Schulinspektionen nicht auf bestehende Indikatoren zurückgegriffen werden. Dies hat jedoch zur Folge, dass deutschlandweit in Bezug auf die Art und Weise der politischen Mitbestimmung unterschiedliche Anforderungen an Schulen formuliert werden. Wenn die Bundesländer zukünftig auf das gleiche Set an Indikatoren zurückgreifen würden, könnte dies dazu führen, dass dieses Indikatorenset an Reichweite gewinnt und somit sich der Grad der Institutionalisierung erhöhen würde. Es gebe demnach theoretisch das Potenzial, dass sich damit Standards politischer Partizipation etablieren.

Des Weiteren ist zu diskutieren, welche Implikationen mit der expliziten Benennung von Differenzlinien einhergehen. So wird in der Inklusionsforschung die Verwendung von Differenzlinien unter anderem kritisch gesehen, wenn diese dazu führen, dass bestehende stigmatisierende Zuschreibungen reproduziert werden (z. B. Wocken 2012; siehe auch Moser 2012a; nähere Ausführung → Kapitel 2), sie aber für das pädagogische Setting irrelevant sind. Dagegen kann die Nicht-Aufführung von Differenzlinien zur Folge haben, dass systematische Formen der Diskriminierung bzw. fehlender Teilhabe nicht sichtbar werden: So wird in der Literatur zur Messung von Bildungsungleichheiten argumentiert: „Was nicht gezählt wird, zählt nicht“ (z. B. Ahyoud et al. 2018; Anagnostopoulos et al. 2016, S. 348; Ainscow 2005; siehe auch Katzenbach 2015; nähere Ausführung → Kapitel 4). Um Einschränkungen von Teilhabe oder Formen von Diskriminierung sichtbar zu machen, wäre es daher notwendig, explizit relevante Differenzlinien zu benennen. In diesem Zusammenhang empfehlen Ahyoud et al. (2018) jedoch, auf Messformen zurückzugreifen, die bei bestehenden Formaten des Bildungsmonitorings bislang kaum etabliert sind, z. B. Indikatoren zur Selbstidentifikation. Hierbei werden Schüler_innen beispielsweise befragt, ob sie von anderen als Nicht-Deutsche wahrgenommen werden und ob dies zur Folge hat, dass sie diskriminiert werden. In Bezug auf Behinderung wäre ein möglicher Indikator: „Haben Sie eine Beeinträchtigung, durch die ihre gesellschaftliche Teilhabe behindert wird?“ (Aikins et al. 2018, S. 40). Demnach müsste nicht auf bestehende Definitionen zurückgegriffen werden, die in Bezug auf Diskriminierung z. T. nicht aussagekräftig sind, wie dies bei

der häufig verwendeten Definition von „Migrationshintergrund“ vom Statistischen Bundesamt⁴ der Fall ist. So gibt es beispielsweise viele Personen, die als Nicht-Deutsche wahrgenommen und diskriminiert werden, z. B. aufgrund ihrer Hautfarbe, aber nicht unter diese Definition fallen. In der konkreten Anwendung in den Messinstrumenten der Schulinspektionen würde dies bedeuten, dass die Schüler_innen im Rahmen des Schülerfragebogens aufgefordert werden, sich selbst in Bezug auf bestimmte Attribute wie Gender zuzuordnen. Im nächsten Schritt könnten einzelne Aspekte, z. B. Möglichkeiten der Partizipation, abgefragt werden. Bei der Auswertung könnten dann – bei ausreichender Fallzahl – die Daten miteinander verknüpft werden, um zu untersuchen, ob die Möglichkeiten der Partizipation zwischen verschiedenen Gruppen variieren. Dieses Erhebungsformat und das dazugehörige Auswertungsverfahren wären bei denjenigen Inspektionen, die aktuell Schülerfragebogen anwenden, theoretisch möglich. Diese Option wird jedoch von Seiten der Schulinspektionen nicht genutzt. Durch Formen der Selbstidentifikation könnten demnach Diskriminierung von Teilgruppen sichtbar werden bzw. lässt sich somit auch fehlende Teilhabe von Einzelpersonen abbilden. Selbstidentifikation im Schülerfragebogen könnte also als Ergänzung zu anderen Erhebungsformaten gesehen werden, um Diskriminierung oder fehlende Teilhabe sichtbar zu machen. In diesem Zusammenhang ist jedoch zu berücksichtigen, dass Formen der Selbstidentifikation nur begrenzt bei allen Differenzlinien möglich sind, z. B. bei Leistung. Daher bieten sich andere Erhebungsformate als Ergänzung an, z. B. Differenzierung im Unterricht über den Unterrichtsbeobachtungsbogen oder allgemeine Frage nach Diskriminierungserfahrungen im Schülerfragebogen.

Schulinspektionen im Vergleich zu anderen Evaluationen: Aus der empirischen Analyse lässt sich die These formulieren, dass die hier untersuchten Messinstrumente der Schulinspektionen sich von anderen datengestützten Evaluationen im Bildungsbereich in Deutschland sowie im Ausland, z. B. in England und in den Niederlanden, unterscheiden. Wie bereits dargestellt (→ Kapitel 3) wurde die Einführung der Schulinspektionen im Fall von Deutschland als Ergänzung zu anderen empirischen Erhebungen wie Lernstandserhebungen gesehen. Während Lernstandserhebungen wie der IQB-Ländervergleich den Fokus auf die Outputebene legen, zielen die Schulinspektionen primär darauf ab, Daten auf der Prozessebene zu erheben. Der Fokus auf die Prozessebene lässt sich auch anhand der hier durchgeführten Analyse der Messinstrumente zu Inklusion bestätigen. Damit unterscheiden sich die deutschen Schulinspektionen von den Verfahren ausländischer Schulinspektionen: So weisen Ehren und Scheerens im Rahmen einer international vergleichenden Analyse von Inspektionsverfahren nach, dass „das Hauptaugenmerk bei Schulinspektionen immer mehr auf dem Output von Schulen liegt“ (vgl. Ehren und Scheerens 2015, S. 235) und dabei insbesondere auf der Messung von Lernfortschritten. Der Fokus auf die Prozessebene durch die deutschen Schulinspektionen ist insofern positiv zu bewerten, als es nach Biesta (2011) in pädagogischen Settings nicht nur relevant ist, ob ein Lernziel erreicht wird, sondern auch welcher Weg eingeschlagen wird, um dieses zu erreichen.

Neben dem Fokus auf die Prozessebene stellt sich die Frage, inwiefern das Verfahren der deutschen Schulinspektionen noch in Bezug auf andere Aspekte eine Alternative zu bestehenden datengestützten Evaluationen darstellt. Für einen Vergleich bieten sich hierbei Lernstandserhebungen wie PISA, TIMSS, IGLU sowie die IQB-Vergleichsstudie an. Diese Untersuchungen

⁴ Definition von „Migrationshintergrund“ nach dem Statistischen Bundesamt: „Eine Person hat einen Migrationshintergrund, wenn sie selbst oder mindestens ein Elternteil nicht mit deutscher Staatsangehörigkeit geboren wurde“ (Statistisches Bundesamt 2018).

werden im Fall von Deutschland auf verschiedenen Ebenen des Bildungssystems dafür genutzt, um Qualitätsurteile von Bildungseinrichtungen zu treffen und zum Teil als Grundlage für bildungspolitische Entscheidungen. Zugleich stehen diese jedoch in der Kritik, weil sie sich auf die Qualifizierungsfunktion von Schule beschränken (Meyer und Zahedi 2014). Indem im Rahmen der empirischen Analyse dieser Arbeit auf ein Verständnis von Inklusion zurückgegriffen wurde, welches verschiedene Funktionen von Schule in den Blick nimmt, wird deutlich, dass Schulinspektionen in Deutschland sich nicht nur auf die Qualifizierungsfunktion konzentrieren, sondern weitere Funktionen, z. B. die Enkulturationsfunktion von Schule (vgl. Fend 2006), in den Blick nehmen. Darüber hinaus unterscheiden sich die deutschen Schulinspektionen von vielen anderen in Deutschland praktizierten Evaluationsverfahren, die auf empirischen Datenerhebungen basieren, darin, dass sie nicht nur auf quantitative Formen der Datenerhebung zurückgreifen und dass sie keine Rankings veröffentlichen. Demnach erhalten die Schulen in den Inspektionsberichten nicht nur numerische Werte als Rückmeldung, sondern Fließtexte, welche es ermöglichen, die Ergebnisse der Inspektionen zu kontextualisieren. Im Fall von Berlin erhalten die Schulen auch die Möglichkeit die Ergebnisse zu kommentieren. Ausgehend von den oben beschriebenen negativen Effekten, die nach Espeland und Sauder (2007) durch die Beschränkung auf numerische Werte und deren Präsentation in Form von Rankings eintreten können, kann das Vorgehen der Schulinspektionen in Deutschland positiv gesehen werden. Jedoch kann die Tatsache, dass Schulinspektionen nicht nur auf numerische Werte zurückgreifen, möglicherweise auch eine Erklärung dafür sein, dass die Wirkung von Schulinspektionen eher niedrig eingestuft wird: So vereinfachen numerische Werte den Vergleich zwischen Schulen, weil sie Visualisierungen und Rankings ermöglichen (vgl. ebd.). Des Weiteren führen sie dazu, dass Qualitätsbeurteilungen ohne viel Zeitaufwand rezipiert werden können. So sind numerische Werte einfacher zu kommunizieren (vgl. Bloem 2016). In diesem Zusammenhang stellt sich jedoch die Frage, inwiefern ein stärkerer Wettbewerb zwischen Schulen, der vermutlich durch den Fokus auf numerische Werte entstehen würde, bildungspolitisch gewollt bzw. sinnvoll ist.

Abschließend lässt sich festhalten, dass Schulinspektionen durch ihren Fokus auf die Prozessebene, durch die Verwendung von qualitativen und quantitativen Formen der Datenerhebung und durch die Thematisierung von verschiedenen Funktionen von Schule als Alternative zu anderen Evaluationsformaten im Bildungsbereich gesehen werden können. In Bezug auf Forschungsdesiderata lässt sich festhalten, dass es bislang an Untersuchungen fehlt, die sich systematisch damit auseinandersetzen, wie Qualität bei anderen Evaluationen im Bildungsbereich, z. B. Deutscher Schulpreis oder Studien zur Kitaqualität, gemessen wird. Hierbei würde es sich insbesondere anbieten, im Rahmen von Vergleichen zu untersuchen, inwiefern bei verschiedenen Evaluationen unterschiedliche Anforderungen an die gleiche Einrichtung, z. B. Kindertagesstätten oder Schulen, formuliert werden und auf welche Qualitätsverständnisse diese zurückzuführen sind. Eine Ausnahme stellt in diesem Zusammenhang die PISA-Studie dar. Das dahinterliegende Verständnis von Schulqualität wurde sowohl in wissenschaftlichen als auch medialen Diskursen breit diskutiert (z. B. Meyer und Zahedi 2014).

9.1.2 Verständnisse von Inklusion in den Instrumenten

Im Rahmen der vorliegenden Studie wurde in einem zweiten Schritt analysiert, welches Verständnis von Inklusion in den Instrumenten sichtbar wurde. Hierfür wurden nur diejenigen Instrumente analysiert, die eindeutig als „inklusionsorientierte“ Instrumente identifiziert wer-

den konnten. Es ist hierbei zu berücksichtigen, dass es vier Bundesländer gab, in denen die Schulinspektionen innerhalb ihres Instrumentensets keine Messinstrumente zur Umsetzung von Inklusion gekennzeichnet hatten: Brandenburg, Hessen, Saarland und Sachsen-Anhalt. Sie werden daher in diesem Teil der Analyse nicht berücksichtigt.

Kurzzusammenfassung der Ergebnisse aus der Analyse: In der Analyse kristallisierten sich vier Typen heraus (nähere Ausführung → Abschnitt 7.2): Diese unterscheiden sich in ihrem Inklusionsverständnis (enger vs. weiter Inklusionsbegriff), in den inhaltlichen Anforderungen, die sich aus der Umsetzung von Inklusion ergeben, im Anwendungsbereich der Indikatoren (alle Schulen vs. bestimmte Schulen) sowie in der Integration in das bestehende Instrumentenset (Querschnitt oder Add-on). Allgemein lässt sich festhalten, dass die meisten Bundesländer bei den betrachteten Instrumenten den Fokus auf die Differenzlinie „Behinderung“ und die Qualifizierungsfunktion von Schule legen.

Unterschiede in der Deutung von Inklusion: Da sich die Instrumente der Schulinspektionen in den einzelnen Bundesländer in ihren Deutungen von Inklusion bzw. in deren Operationalisierung unterscheiden, bestätigt die hier durchgeführte Analyse die Ergebnisse anderer empirischer Studien, die nachweisen konnten, dass es auf Bundesländerebene Unterschiede in der Implementation von Inklusion gibt (z. B. Blanck et al. 2013; Gasterstädt 2016; Bertelsmann-Stiftung 2014). Somit gibt es auch auf der Ebene der Bildungsadministration keinen Konsens darüber, was unter Inklusion verstanden wird (vgl. Moser 2012b; Grosche 2015). Des Weiteren stehen die Ergebnisse der hier durchgeführten Analyse im Widerspruch zu Aussagen aus den Interviews, wonach die Bundesländer bei der Messung von Schulqualität auf die gleichen Instrumente zurückgreifen. Dies lässt sich mit Blick auf einen Teilaspekt von Schulqualität, der Umsetzung von Inklusion, nicht bestätigen. Um hierzu weitere Aussagen treffen zu können, bedarf es jedoch empirischer Untersuchungen, die die Instrumente der Schulinspektionen bezüglich weiterer Aspekte von Schulqualität miteinander vergleichen. Hierzu fehlt es bislang an Studien.

Wie bereits im Inklusionskapitel dargestellt, gibt es für Deutschland bislang kaum Studien, die auf Bundeslandebene die Umsetzung von Inklusion erfassen und hierbei verschiedene Ebenen (Input-, Prozess- und Outputebene) sowie verschiedene inhaltliche Aspekte und Funktionen von Schule in den Blick nehmen. Durch die Studien von Klemm (vgl. Bertelsmann-Stiftung 2014), die bundesländerübergreifend Aussagen zulassen, wird lediglich abgebildet, inwiefern es auf der Systemebene bei der Platzierung von Schüler_innen mit Behinderung in der Regelschule Unterschiede zwischen den Bundesländern gibt. So lassen sich mit diesen Daten Aussagen darüber treffen, wie viele Schüler_innen mit sonderpädagogischem Förderbedarf an Regel- und Förderschulen beschult werden. Da diese Studien nur eine Personengruppe, Schüler_innen mit sonderpädagogischem Förderbedarf, in den Blick nehmen und primär den Zugang zu Schule betrachten, können diese Daten nur begrenzt Aussagen darüber treffen, inwiefern sich die Bundesländer bei der Implementation von Inklusion auf Einzelschulebene, z. B. Unterrichtsebene, unterscheiden. Demnach können die Ergebnisse dieser Arbeit in Hinblick auf Bundesländerunterschiede nur begrenzt mit anderen Studien verglichen werden.

Es lässt sich jedoch festhalten, dass die Ergebnisse aus diesem Teil der Analyse Untersuchungen bestätigen, die sich generell mit der Implementation von politischen Vorgaben auseinandersetzen (z. B. Spillane et al. 2002; Fend 2008; Amrhein 2011): So bestätigt auch diese Studie, dass die gleiche politische Vorgabe – in diesem Fall Inklusion – in Abhängigkeit vom jeweiligen Kontext

unterschiedlich gedeutet wird und demnach davon auszugehen ist, dass politische Vorgaben selten in der Praxis so implementiert werden, wie sie intendiert waren.

Des Weiteren stellt sich die Frage, welche Folgen die unterschiedlichen Deutungen von Inklusion bzw. deren Operationalisierung in den Instrumenten für die Schulen haben: So lässt sich festhalten, dass es in den jeweiligen Bundesländern variiert, welche Anforderungen in Bezug auf Inklusion an die Schulen formuliert werden. Des Weiteren gibt es auf Bundeslandebene Unterschiede darin, welche Schulen die jeweiligen Anforderungen erfüllen müssen und inwiefern die Schulen selbst entscheiden können, ob sie diese Anforderungen erfüllen wollen oder nicht. Dies gibt auch Hinweise darauf, welche Relevanz der Umsetzung von Inklusion im jeweiligen Bundesland zugeschrieben wird.

In diesem Zusammenhang sind auch die Bundesländer zu nennen, bei denen nicht erkennbar ist, dass sie Instrumente zu Inklusion in das bestehende Instrumentenset integriert haben (BB; HE; SL; ST). Hierfür gibt es zwei mögliche Erklärungen: Zum einen kann es sein, dass sich die Schulinspektionen in den jeweiligen Bundesländern dafür entschieden haben, dass die Implementation von Inklusion kein Qualitätsmerkmal sei, das von Seiten der Schulinspektion überprüft werden müsste, z. B. weil die Implementation auf Schulebene sich in der Anfangsphase befindet oder weil sie die Umsetzung von Inklusion im Hinblick auf die Messung von Schulqualität nicht für relevant erachten. Zum anderen gibt es auch den Fall, dass bestimmte Indikatoren bzw. Instrumente bewusst nicht als inklusionsorientiert gekennzeichnet wurden, aber das bestehende Instrumentenset solche enthält. In den Interviews wurde dieses Vorgehen als bewusst getroffene Strategie beschrieben: Diese sei notwendig, wenn andere Akteure im Bildungsbereich, z. B. das jeweilige Kultusministerium, ein Verständnis von Inklusion vertraten, welches von ihrem Verständnis abwich oder wenn sie der Meinung waren, dass die Umsetzung von Inklusion von Seiten der Schule als große Herausforderung wahrgenommen wird und sie daher aus strategischen Gründen den Begriff vermeiden wollten. (Aus Datenschutzgründen können diese Aussagen aus den Interviews nicht konkreten Bundesländern zugeordnet werden.)

Instrumente zur Messung von Inklusion in Abhängigkeit von den jeweiligen Funktionen von Schulinspektionen:

Schulinspektionen werden im Bildungssystem verschiedene Funktionen zugeschrieben, z. B. Kontroll- und Impulsfunktion (nähere Ausführung → Kapitel 3). In Bezug auf die explizit gekennzeichneten Instrumente zu Inklusion lässt sich zum Teil ein Zusammenhang zu einzelnen Funktionen herstellen: So scheint die Kontrollfunktion zu dominieren, wenn Instrumente zu Inklusion für die Schulen verpflichtend sind und klare Anforderungen gegenüber Schulen formuliert werden. Dagegen spiegelt sich die Impulsfunktion in Instrumenten wider, die für die Schulen optional sind oder wenn von Seiten der Schulinspektion nur abgefragt wird, ob die Schulen sich schon einmal mit Inklusion auseinandergesetzt haben.

Die jeweiligen Ausrichtungen bei der Entwicklung der Instrumente haben für die Schulen unterschiedliche Implikationen: Wenn die Instrumente zu Inklusion die Impulsfunktion von Schulinspektionen widerspiegeln, wird gegenüber Schulen nicht die Erwartung formuliert, dass sie die jeweiligen Anforderungen erfüllen müssen, sondern sie erhalten Ideen und Hinweise, welche Schritte für die Umsetzung von Inklusion notwendig sind, und erfahren, wo sie sich im Prozess der Umsetzung befinden. Diese Form der Operationalisierung (Typ C: Bund-Länder-Inspektion und Berlin; → Abschnitt 7.2) berücksichtigt, dass die Umsetzung von Inklusion in der Praxis nicht erzwungen werden kann und Schulen können selbst auswählen, ob bei ihnen die Umsetzung von Inklusion evaluiert werden soll. Demnach haben Schulen die Möglichkeit,

sich mit der Umsetzung von Inklusion auseinanderzusetzen und selbst zu bestimmen, wann sie hierzu Rückmeldung von Seiten der Schulinspektion erhalten möchten. Hierbei geht jedoch der Aspekt verloren, dass Schulen gesetzlich zur Umsetzung von Art. 24 BRK verpflichtet sind (vgl. Wrase 2015). Aus dieser Perspektive ist es daher kritisch zu sehen, wenn Schulen frei wählen können, ob sie zur Umsetzung von Inklusion evaluiert werden möchten.

Dagegen lässt sich in anderen Bundesländern bei den Instrumenten zu Inklusion ein Bezug zur Kontrollfunktion von Schulinspektionen herstellen: Dies ist beispielsweise der Fall, wenn Schulinspektionen mit ihren Instrumenten sicherstellen wollen, dass Schüler_innen mit diagnostizierter Behinderung im schulischen Kontext ausreichend Unterstützung erhalten und die dazugehörigen Indikatoren für die jeweiligen Schulen verpflichtend sind (Typ B: NRW; NI; BW; nähere Ausführung → Abschnitt 7.2).

In diesem Fall verstehen sich die Schulinspektionen in Bezug auf Inklusion als kontrollierende Instanz. Hierbei stellt sich jedoch die Frage, wie die jeweiligen Schulinspektionen die konkreten Anforderungen, die sich aus den jeweiligen Instrumenten ergeben, legitimieren können. Im Rahmen der Interviews distanzieren sich die meisten Instrumenteentwickler_innen von der Aufgabe, gesetzliche Vorgaben zu überprüfen. Des Weiteren machen sie deutlich, dass die gesetzlichen Vorgaben bei der Formulierung der Instrumente eine geringe Rolle gespielt haben. Stattdessen wird beispielsweise darauf verwiesen, dass mit den Anforderungen, die in den Instrumenten formuliert werden, der Standard sichergestellt werden soll, der bislang an Förderschulen galt bzw. bringen die Instrumenteentwickler_innen Erfahrungen aus der eigenen Praxis ein. Es gibt jedoch keine rechtliche Legitimation derjenigen Instrumente, die sich auf Inklusion beziehen. Dies ist insofern kritisch zu sehen, als – basierend auf der Kontrollfunktion – den Schulen signalisiert wird, dass sie diese Anforderungen erfüllen müssen, wenn sie Schüler_innen mit Behinderung unterrichten, aber es fehlt an Transparenz darüber, woher die jeweiligen Anforderungen kommen. Gleichwohl wäre ein rechtliche Legitimation z. T. möglich, da Inklusion gesetzlich verankert ist und sich daraus bestimmte Anforderungen an die Einzelschule ergeben (vgl. Piezunka et al. 2018). In diesem Zusammenhang ist zu beachten, dass Schulinspektionen nur begrenzt die Möglichkeit haben Sanktionen auszusprechen, wenn sie im Rahmen der Evaluation bemerken, dass Gesetzesvorgaben wie Inklusion nicht richtig umgesetzt werden.

Daran anknüpfend ist kritisch zu hinterfragen, welche Implikationen damit einhergehen, wenn im Rahmen der Kontrollfunktion die Umsetzung von Inklusion nur als Rechtsanspruch kontrolliert werden würde: So argumentiert unter anderem Felder (2012), dass Inklusion als Rechtsanspruch allein nicht umgesetzt werden kann. Dies basiert auf einem Verständnis von Inklusion, wonach deren Umsetzung impliziert, dass im schulischen Alltag bestimmte Haltungen gelebt werden und sich bestimmte Kulturen etablieren. Demnach würde die Überprüfung von konkreten Maßnahmen, die gesetzlich in Art. 24 BRK verankert sind, z. B. Platzierung in der Regelschule oder aktive Einbeziehung im Unterricht, zu kurz greifen, da weitere Gelingensbedingungen, z. B. Einstellungen der pädagogischen Fachkräfte oder Willkommenskultur, dadurch nicht abgedeckt werden. Diese Aspekte lassen sich jedoch nur schwer im Rahmen der Kontrollfunktion, d. h. mit verpflichtenden Instrumenten erfassen: So lassen sich bestimmte Haltungen nicht erzwingen bzw. würde deren Abfrage bei der Inspektion vermutlich zu sozial erwünschten Antworten führen.

Demnach würde sich in der Praxis eine Kombination aus Kontroll- und Impulsfunktion anbieten: Im Rahmen der Kontrollfunktion könnten Schulinspektionen die Einhaltung von vorhandenen Gesetzesvorgaben überprüfen. Jene Instrumente wären dann verpflichtend für

Schulen. Dabei ist es sinnvoll auf bestehende Gesetzesvorgaben zu verweisen, um die konkreten Instrumente zu legitimieren. In Bezug auf die Überprüfung von Gesetzesvorgaben müsste jedoch auf der Ebene der Bildungsadministration eine Auseinandersetzung darüber stattfinden, ob dies eine Funktion von Schulinspektionen sein soll und welche Folgen die Nichteinhaltung von gesetzlichen Vorgaben hat. Im Hinblick auf die Impulsfunktion von Schulinspektionen sollten im Instrumentenset Indikatoren aufgeführt werden, die mögliche Gelingensbedingungen für die Umsetzung von Inklusion beschreiben.

Des Weiteren stellt sich auch die Frage, inwiefern die Evaluation von Seiten der Schulinspektion auf der Metaebene den Anspruch haben sollte, inklusionsorientiert zu sein. Anders formuliert: Inwiefern fühlen sich schulische Vertreter, wie Lehrkräfte oder Schulleitungen, wahrgenommen und wertgeschätzt, wenn sie evaluiert werden bzw. inwiefern haben sie die Möglichkeit, am Evaluationsprozess teilzuhaben und diesen mitzugestalten. So impliziert die Kontrollfunktion von Schulinspektion eine klare Hierarchie zwischen Schulinspektion und der Schule, wohingegen die Impulsfunktion es ermöglicht, dass beide Akteure sich auf Augenhöhe begegnen (vgl. Dietrich und Lambrecht 2012). Im Rahmen der Arbeit lag der Fokus auf der Analyse der Instrumente sowie der Interviews mit Instrumenteentwickler_innen. Um Aussagen darüber treffen zu können, ob das Handeln von Schulinspektor_innen inklusionsorientiert ist bzw. überhaupt sein kann, wären weiterführende Studien notwendig.

9.1.3 Schlussfolgerungen aus dem zweistufigen Analyseverfahren

Indem die Berücksichtigung von Inklusion in den Instrumenten durch ein zweistufiges Verfahren analysiert wurde, lassen sich für die Implementation von Inklusion als politische Vorgabe verschiedene Schlussfolgerungen ziehen, die im Folgenden dargestellt werden.

„Es steckt mehr Inklusion drin als draufsteht“ : Durch die erste Analyse (→ Abschnitt 7.1) wurde deutlich, dass in den Messinstrumenten Indikatoren zur Messung von Schulqualität vorhanden sind, die Anknüpfungspunkte zu dem in dieser Arbeit verwendeten Verständnis von Inklusion bieten. Diese bleiben jedoch häufig unter den Anforderungen, die sich aus dem hier verwendeten Inklusionsverständnis ergeben, da sie zu wenig die Situation von einzelnen Schüler_innen in den Blick nehmen, obwohl dies aufgrund der verwendeten Erhebungsverfahren möglich wäre.

Im Vergleich zu den Instrumenten, die explizit als „Instrumente zu Inklusion“ gekennzeichnet sind, zeigt sich jedoch, dass in dem gesamten Instrumentenset der Schulinspektionen häufig mehr inklusionsorientierte Aspekte enthalten sind, als dies von Seiten der Inspektionen gekennzeichnet wurde. So konzentrieren sich die als „inklusionsorientiert“ gekennzeichneten Instrumente – wenn sie überhaupt vorhanden sind – in den meisten Bundesländern auf Schüler_innen mit Behinderung und es werden stets nur Teilaspekte von Teilhabe und Diskriminierung operationalisiert. So liegt der Fokus häufig auf der Qualifizierungsfunktion von Schule. Dagegen bieten die anderen Instrumente, die nicht explizit als „inklusionsorientiert“ gekennzeichnet sind, weitere Überschneidungen mit dem hier verwendeten Verständnis von Inklusion. So werden neben der Qualifizierungsfunktion noch weitere Funktionen in den Blick genommen und hierbei werden – neben Behinderung – weitere Differenzlinien aufgeführt. Demnach enthalten auch die Bundesländer, die über keine Instrumente verfügen, die explizit als „inklusionsorientiert“ gekennzeichnet sind, Instrumente, die es erlauben, die Umsetzung von Teilaspekten von Inklusion zu messen.

Stattgefundene Implementation von Inklusion als politische Vorgabe: Nach der Betrachtung davon, wie Inklusion in den Messinstrumenten operationalisiert wurde, stellt sich die Frage, inwiefern die analysierten Instrumente ein Ergebnis davon sind, dass Inklusion als politische Vorgabe in den Messinstrumenten der Schulinspektionen implementiert wurde. Diese Frage lässt sich durch die empirische Analyse, die im Rahmen der Arbeit stattgefunden hat, nicht abschließend beantworten. Hierfür wäre es notwendig gewesen, das ganze Instrumentenset zu zwei unterschiedlichen Zeitpunkten, z. B. vor und nach der Ratifizierung der BRK, zu analysieren. Jedoch liegt die Schlussfolgerung nahe, dass die nicht als „inklusionsorientiert“ gekennzeichneten Instrumente unabhängig von der Ratifizierung der BRK entwickelt wurden, da beispielsweise die Forderung nach „innerer Differenzierung“ als solche bereits seit längerem im pädagogischen Diskurs – unabhängig von Inklusion – diskutiert wird. So weisen Trautmann und Wischer nach, dass viele Aspekte von „innerer Differenzierung“ bereits in den 70er Jahren diskutiert wurden und es eine hohe Übereinstimmung zwischen den theoretischen Überlegungen von damals und heute gibt (vgl. Trautmann und Wischer 2009). Demnach lässt sich lediglich festhalten, dass bei den Schulinspektionen Instrumente vorhanden sind, die normative Vorstellungen von guter Schule repräsentieren, welche Anknüpfungspunkte zu dem in dieser Arbeit verwendeten Verständnis von Inklusion ermöglichen. Es lässt sich daraus jedoch nicht schlussfolgern, dass diese als Reaktion auf die Ratifizierung der BRK entstanden sind. In Bezug auf die Instrumente, die als „inklusionsorientiert“ von Seiten der Inspektionen gekennzeichnet wurden, lässt sich eine andere Schlussfolgerung ziehen: So zeigte sich in der Studie von Degenhardt, der die Instrumente der Schulinspektionen in 2008 untersuchte, dass der Begriff Inklusion „in keinem der analysierten Handbücher vertreten“ (Degenhardt 2008, S. 21) war. Die Analyse, die im Rahmen dieser Arbeit durchgeführt wurde, zeigt bei vielen Bundesländern ein anderes Ergebnis. So gibt es in acht der zwölf analysierten Bundesländer Instrumente zu Inklusion bzw. zum gemeinsamen Unterricht von Schüler_innen mit und ohne Behinderung. Demnach liegt es nahe, dass das Vorhandensein von Instrumenten, die als „inklusionsorientiert“ gekennzeichnet sind, eine Folge davon ist, dass Deutschland die BRK ratifiziert hat bzw. dass Inklusion als politische Vorgabe Eingang in den öffentlichen Diskurs gefunden hat.

9.1.4 Deutungen von Seiten der Instrumenteentwickler_innen

Im Folgenden wird dargestellt, wie Inklusion auf individueller Ebene durch die Instrumenteentwickler_innen gedeutet wurde. Diese Deutungen sind das Ergebnis von individuellen Sensemaking-Prozessen und dienen somit als Grundlage, um im nächsten Schritt die Rahmenbedingungen, die den Deutungsprozess geprägt haben, nachvollziehen zu können.

Kurz Zusammenfassung der Ergebnisse aus der Analyse: Durch die Analyse der Interviews ließen sich drei Deutungen von Inklusion herausarbeiten:

- Typ I: Inklusion wird als ein Anspruch auf Unterstützung angesehen, der für Schüler_innen mit Behinderung (= Zielgruppe) an der Regelschule besteht. Schulinspektionen möchten sicherstellen, dass dieser Anspruch von Seiten der Schulen eingelöst wird.
- Typ II: Inklusion – verstanden als gemeinsamer Unterricht von Schüler_innen mit und ohne Behinderung – wird als eine Herausforderung für Schulen gesehen, die sich aus den rechtlichen Vorgaben ergibt. Die Inspektion will mehr darüber erfahren, ob sich die Schulen mit dem Thema bereits auseinandersetzen.

- Typ III: Unter Inklusion wird verstanden, wenn es Schulen gelingt, auf die Heterogenität der Schülerschaft einzugehen und die Schulinspektion überprüft die Implementation von Inklusion, weil dies ein Teilaspekt von Schulqualität darstellt.

Die jeweiligen Deutungen unterscheiden sich unter anderem darin, was unter Inklusion verstanden wird und warum die Schulinspektion die Implementation von Inklusion auf schulischer Ebene überprüfen sollte. Ausgehend von den jeweiligen Deutungen von Inklusion ergeben sich unterschiedliche Vorstellungen davon, wie die Umsetzung von Inklusion in den Messinstrumenten operationalisiert werden sollte. So unterscheiden sich die jeweiligen Vorstellungen bei der inhaltlichen Ausgestaltung, bei der Integration in das bestehende Instrumentenset der Schulinspektionen sowie beim Grad der Anforderungen. Hierbei ist jedoch zu berücksichtigen, dass diese Operationalisierungen, die in den Interviews genannt wurden, sich nicht immer in den tatsächlichen Instrumentensets der jeweiligen Bundesländer wiederfinden. Dies hat verschiedene Gründe, z. B. dass sich Einzelpersonen mit ihren Deutungen von Inklusion innerhalb ihres Teams bzw. gegenüber hierarchiehöheren Akteuren nicht durchsetzen konnten (nähere Ausführung → Abschnitt 8.4). Dieser Aspekt ist auch zu beachten, wenn es im Folgenden um die Entwicklung der Instrumente geht.

Bezüge zu anderen Deutungen von Inklusion auf Individualebene : Bei den Deutungen durch die Instrumenteentwickler_innen zeigte sich ein ähnliches Phänomen, dass bereits bei der empirischen Analyse der Instrumente sichtbar wurde und bestehende Studien aus der Inklusionsforschung bestätigt: Es gibt keine einheitliche Deutung von Inklusion im schulischen Kontext (→ Abschnitt 9.1.2).

In Bezug auf diesen empirischen Teil bietet es sich an, Bezüge zu Studien herzustellen, die sich damit auseinandergesetzt haben, wie schulische Inklusion von Einzelpersonen gedeutet wurde. So hat sich DeMatthews (2012) mit der Frage auseinandergesetzt, wie verschiedene Schulleitungen eines US-amerikanischen Bezirks bei der Deutung von Inklusion vorgegangen sind. Mithilfe des Sensemaking-Ansatzes untersuchte er unter anderem, welche "cues", d. h. einzelne Aspekte der politischen Vorgabe, von Seiten der Schulleitungen in den Blick genommen wurden, um die Vorgabe als solche gedanklich bearbeitbar zu machen. Hierbei zeigte sich bei DeMatthews, dass der Fokus auf der Frage lag, wie Schüler_innen mit Behinderung in der Regelschule zu platzieren sind:

„Some principals concluded that they needed to move all or most of their students into the general education classroom while other principals concluded that they had to follow specific policies and procedures to maintain segregated environments for certain students“ (ebd., S. 211).

Dieser Aspekt spielte für die Instrumenteentwickler_innen der Schulinspektionen keine Rolle. Dagegen zeigte sich in der empirischen Analyse, dass die Zielgruppe von Inklusion ein wichtiger "cue" für die Instrumenteentwickler_innen war. So setzten sie sich intensiv mit der Frage auseinander, wer die Zielgruppe von Inklusion ist. Hier lassen sich Anknüpfungspunkte zu Studien herstellen, die sich mit der Existenz von unterschiedlichen Inklusionsverständnissen auseinandersetzen, z.B. von Göransson und Nilholm (2014); Kruse und Dederich (2018); Ainscow et al. (2006); Hinz (2002) sowie Piezunka et al. (2017). So haben auch diese Studien den Aspekt der Zielgruppe für bedeutsam identifiziert, um zwischen verschiedenen Inklusionsverständnissen zu unterscheiden.

9.2 Entwicklung von Indikatoren zu Inklusion

Durch die empirische Analyse der Interviews wurde sichtbar, welche Rahmenbedingungen das Sensemaking zu Inklusion geprägt haben und auf welche Handlungsmuster bei der Entwicklung von Instrumenten zurückgegriffen wird. Bislang gibt es wenige Studien, die sich damit auseinandersetzen, wie unterschiedliche Formen der Deutung in der Praxis zu erklären sind, z. B. verweisen Blanck et al. (2013) unter anderem auf Pfadabhängigkeiten, die sich aus historisch gewachsenen Strukturen der jeweiligen Bundesländer ergeben. Aufgrund fehlender Forschung können im Folgenden daher nur begrenzt Bezüge zu Studien hergestellt werden, die sich mit ähnlichen Fragestellungen auseinandergesetzt haben (zum Forschungsstand → Kapitel 4 und → Kapitel 2).

9.2.1 Kurzzusammenfassung der Ergebnisse aus der Analyse

Bei den Interviews wird deutlich, welche Rahmenbedingungen das Sensemaking zu Inklusion bzw. zu den dazugehörigen Instrumenten geprägt haben. Im Folgenden wird auf die einzelnen Elemente (persönliche Attribute, situativer Kontext sowie politische Vorgabe), die von Spillane et al. (2002) eingeführt wurden, näher eingegangen. Hierbei ist zu berücksichtigen, dass manche Rahmenbedingungen auf die Deutung von Inklusion prägend wirkten und andere sich stärker auf die sich daraus ergebenden Instrumente zur Messung von Inklusion auswirkten. So zeigte sich beispielsweise in der Analyse, dass einzelne Personen zwar das gleiche Verständnis von Inklusion haben, aber unterschiedliche Positionen vertreten, wie die Umsetzung von Inklusion in den Messinstrumenten der Schulinspektionen operationalisiert werden sollte.

Bei der Dimension „**persönliche Attribute**“ wird häufig auf die eigene Praxiserfahrung als Lehrkraft verwiesen sowie auf Beobachtungen, welche die Interviewpartner im Rahmen ihrer Tätigkeit als Inspektor_innen gemacht haben. Die Interviewpartner unterscheiden sich unter anderem darin, ob sie Erfahrung im gemeinsamen Unterricht haben, inwiefern sie über Kenntnisse im Bereich der Sonderpädagogik oder zu Inklusion / Integration verfügen, welche Position sie bei der Umsetzung von Inklusion vertreten und was sie unter einer guten Schule verstehen. In Bezug auf Inklusion wird den Personen, die über sonderpädagogische Qualifikationen bzw. Erfahrungen im gemeinsamen Unterricht von Schüler_innen mit und ohne Behinderung verfügen, häufig Expertise zugeschrieben bzw. die Personen mit der jeweiligen Qualifikation begründen ihre eigene Expertise damit.

Als weitere Dimension ist der „**situative Kontext**“ zu nennen: Hierbei spielen eine Vielzahl von Akteuren und Kontextbedingungen eine Rolle. Diese haben unter anderem Einfluss darauf, welche Anforderungen an Schule formuliert werden und wie der Begriff „Inklusion“ in den Instrumenten verwendet wird. Als relevante Akteure sind in diesem Zusammenhang die Kultusministerien, schulische Vertreter sowie die Schulaufsicht zu nennen. So variiert es zwar zwischen den Bundesländern, zu welchem Grad jene Akteure bei der Entwicklung von Instrumenten einbezogen werden. Aber es zeigt sich einzelfallübergreifend, dass die Interessen und Positionen der genannten Akteure bei der Entwicklung der Instrumente angehört bzw. antizipiert wurden. Des Weiteren verwiesen viele Instrumenteentwickler_innen auf den Austausch zwischen den Schulinspektionen der einzelnen Bundesländer in Bezug auf Inklusion. Im Rahmen dieser alljährlichen Treffen⁵ präsentieren zwar die Vertreter der Bundesländer ihre Instrumente zu

⁵ Die Autorin der Arbeit war persönlich bei den Jahrestreffen in Soest (2015), Wiesbaden (2016) und Speyer (2017) dabei. Hierbei wurden keine Daten erhoben, jedoch dienen die erhaltenen Informationen als Hintergrundwissen bei der Analyse der Interviews.

Inklusion, aber eine Harmonisierung, d. h. eine Angleichung der Instrumente zur Erfassung der Umsetzung von Inklusion, wurde bzw. wird nicht angestrebt.

Neben den oben genannten Kontextbedingungen, die dem situativen Kontext zuzuordnen sind, zeigte sich, dass auch entscheidend war, was als primäre Funktion von Schulinspektion angesehen wird. So gab es insbesondere unterschiedliche Positionen zu der Frage, inwiefern die Instrumente zu Inklusion im Sinne der Kontrollfunktion Minimalstandards sein sollten oder im Sinne der Impulsfunktion neue Impulse setzen sollten. Als weitere Rahmenbedingung ist die Forschung zu nennen: So wird in den Interviews darauf verwiesen, dass bei der Entwicklung von Messinstrumente auf den bestehenden Status Quo der Forschung zurückgegriffen wird. Im konkreten Fall, der Entwicklung von Messinstrumenten zur Umsetzung von Inklusion, zeigte sich jedoch, dass die bestehende Forschung nur selten für hilfreich erachtet wurde bzw. diese für die Entwicklung von Instrumenten genutzt wurde.

Als dritte Dimension, die nach dem Sensemaking-Ansatz die Deutung von Inklusion prägen würde, ist die politische Vorgabe – in diesem Fall Inklusion – als solche zu nennen. Dies können beispielsweise konkrete Maßnahmen eines Gesetzes oder Deutungen von politischen Vorgaben sein, die im öffentlichen Diskurs bereits existieren und die von Seiten der Instrumenteentwickler_innen herangezogen werden, wenn sie Inklusion deuten und in ihrem Handlungsbereich umsetzen: Hierbei lässt sich festhalten, dass gesetzliche Vorgaben, wie die BRK sowie relevante Gesetzesvorgaben und Empfehlungen auf Bundeslandebene, dazu geführt haben, dass die Auseinandersetzung mit Inklusion in den Schulinspektionen initiiert wurde.

Auch wenn die Rahmenbedingungen auf individueller Ebene sowie auf Bundeslandebene variieren, lassen sich **drei Handlungsmuster** herausarbeiten, die einzelfallübergreifend für die Instrumenteentwickler_innen relevant waren: Als **erstes Handlungsmuster** zeigte sich, dass die Instrumenteentwickler_innen häufig auf ihre eigenen Erfahrungen – sei es als Lehrkraft oder als Inspektor_in – zurückgriffen, um zu beantworten, durch welche konkreten Indikatoren die Umsetzung von Inklusion gemessen werden kann und welche Anforderungen an Schulen formuliert werden können bzw. sollten. Als **zweites Handlungsmuster** ist das Streben nach Akzeptanz zu nennen: Dies zeigt sich darin, dass bei der Entwicklung von Instrumenten die Positionen von bestimmten Akteuren (schulische Vertreter, Kultusministerien sowie Schulamt) für die Instrumenteentwickler_innen eine wichtige Rolle spielen. Die Instrumenteentwickler_innen gehen davon aus, dass eine starke Einbeziehung dieser Akteure dazu führt, dass das Inspektionsverfahren als solches sowie die Instrumente höhere Akzeptanz erfahren. Als **drittes Handlungsmuster** ist die Berücksichtigung von gesetzlichen Vorgaben zu erwähnen: Wie bereits oben dargestellt, haben diese dazu geführt, dass Prozesse in den einzelnen Bundesländern initiiert wurden, aber bei der Frage danach, welche Anforderungen an Schule formuliert werden und wie dies gemessen werden kann, wurden in den Interviews kaum Bezüge zu gesetzlichen Vorgaben hergestellt: So wurde nur einmal thematisiert, dass die BRK als Text über den Email-Verteiler der jeweiligen Schulinspektion verschickt wurde. Ansonsten sprach kein Interviewpartner an, dass die Gesetzestexte bei der konkreten Entwicklung von Messinstrumenten von Bedeutung gewesen wären. Dies wurde unter anderem damit begründet, dass es nicht die Aufgabe der Schulinspektion sei, die Einhaltung von gesetzlichen Vorgaben zu überprüfen.

9.2.2 Gründe für unterschiedliche Deutungen

Nach der Betrachtung der relevanten Rahmenbedingungen sowie der Handlungsmuster, die sich fallübergreifend zeigten, stellt sich die Frage, **wie es auf Individualebene zu unterschiedli-**

chen Deutungen von Inklusion bzw. auf Bundeslandebene zu voneinander abweichenden Instrumenten zu Inklusion gekommen ist. Hierbei ist zu berücksichtigen, dass im Rahmen des Sensemaking-Ansatzes davon ausgegangen wird, dass die einzelnen Rahmenbedingungen im Deutungsprozess miteinander interagieren und somit nicht als voneinander unabhängige Faktoren gesehen werden können. Des Weiteren spielt auch eine Rolle, wie die einzelnen Rahmenbedingungen von Seiten der Instrumenteentwickler_innen gewichtet werden (nähere Ausführung → Kapitel 5).

Bei der Deutung von Inklusion bzw. den jeweiligen Instrumenten zeigt sich, dass die Rahmenbedingungen bei den individuellen Sensemakern sowie in ihrem jeweiligen situativen Kontext variieren. In Bezug auf den situativen Kontext verweisen die Interviewpartner beispielsweise darauf, dass es zwischen den Bundesländern Unterschiede darin gibt, wie viel Erfahrung die Schulen mit der Umsetzung von Inklusion haben. Des Weiteren gab es bei den einzelnen Sensemakern unterschiedliche Positionen dazu, ob die Umsetzung von Inklusion positiv zu bewerten sei.

In diesem Zusammenhang ist auch zu berücksichtigen, dass einzelfallübergreifend zwar die gleichen Akteure oder die gleichen Attribute von Relevanz waren, dass von diesen jedoch unterschiedliche Impulse ausgingen: Beispielsweise spielten die Interessen der Kultusministerien für viele Sensemaker eine wichtige Rolle, aber die Ministerien variierten in ihren Positionen dahingehend, wie Inklusion in den Instrumenten operationalisiert werden sollte. Auf Individualebene stellte sich heraus, dass ähnliche Attribute von Seiten der Instrumenteentwickler_innen als relevant für den Deutungsprozess angesehen wurden. Dies implizierte aber nicht, dass aus dem gleichen Attribut, z. B. sonderpädagogische Qualifizierung, die gleichen Schlussfolgerungen gezogen wurden. Dies lässt sich in diesem Fall unter anderem dadurch erklären, dass sie in ihrer Rolle als Sonderpädagog_innen in unterschiedlichen beruflichen Settings, z. B. Förderschule oder Integrationsschulen, tätig waren. Des Weiteren variierte auch, **welche Relevanz den einzelnen Rahmenbedingungen auf Individualebene zugeschrieben wurde**, z. B. wie stark die Interessen der schulischen Vertreter angehört werden. Darüber hinaus ist auch zu berücksichtigen, dass die einzelnen Sensemaker sich in einem kollektiven Aushandlungsprozess mit anderen Instrumenteentwickler_innen im jeweiligen Bundesland befanden. In vielen Bundesländern gab es innerhalb der Schulinspektion übereinstimmende Positionen dazu, wie Inklusion gedeutet werden sollte. Gleichwohl gab es in einzelnen Schulinspektionen auch die Situation, dass Kompromisse geschlossen werden mussten. Mit der Folge, dass die im Rahmen dieser Arbeit interviewten Sensemaker sich innerhalb ihrer Schulinspektion mit ihren Deutungen manchmal nur zum Teil oder kaum durchsetzen konnten.

Abschließend lässt sich in Bezug auf die jeweiligen Deutungen von Inklusion bzw. Instrumenten zu Inklusion die Annahme des Sensemaking-Ansatzes insofern bestätigen, als sich die Unterschiede zwischen den Bundesländern nicht durch einfache Ursache-Wirkungs-Modelle erklären lassen. Stattdessen zeigte sich in der Analyse, dass auf Individualebene sowie im kollektiven Aushandlungsprozess auf Bundeslandebene variierte, welche Rahmenbedingungen relevant waren, welche Impulse von diesen ausgingen und wie diese von Seiten der Instrumenteentwickler_innen gewichtet wurden. Dabei sind es zum Teil unterschiedliche Konstellationen, die zu den gleichen Ergebnissen führten bzw. ähnliche Konstellationen, die zu unterschiedlichen Deutungen führten, weil beispielsweise einzelne Rahmenbedingungen bei den jeweiligen Personen abwichen oder diese zum Teil anders gewichtet wurden.

In diesem Zusammenhang bietet sich als Analogie das Backen eines Kuchens an: Auf der Grundlage des gleichen Kuchenrezepts (~politische Vorgabe) können sehr unterschiedliche Ergebnisse

(-Deutungen von Inklusion) herauskommen, wenn z. B. eine Zutat (-Rahmenbedingungen) ergänzt wird, ein andere Marke (-variierende Impulse) verwendet wird oder bei einer Zutat die doppelte Menge (~ unterschiedliche Gewichtung der Rahmenbedingungen) hinzugefügt wird. Des Weiteren können Kuchen auch sehr ähnlich schmecken, obwohl unterschiedliche Zutaten verwendet wurden. Die Analogie wirkt auf den ersten Blick banal, macht aber die Zusammenhänge zwischen den einzelnen Rahmenbedingungen und der davon ausgehenden Impulse sowie den unterschiedlichen Deutungen bzw. Handlungsmustern, die das Ergebnis dieser Interaktionen sind, deutlich. Darüber hinaus wird durch diese Analogie deutlich, dass die stattfindenden Mechanismen beim Sensemaking-Ansatz – im Gegensatz zu anderen Theorien – in der Analyse sich nicht klar voneinander trennen lassen.

Für die Praxis, z. B. die Bildungspolitik, haben jene Ergebnisse zur Folge, dass diese, wenn sie den Entwicklungsprozess der Instrumente beeinflussen möchten, zwar versuchen können, auf einzelne Rahmenbedingungen Einfluss zu nehmen, z. B. auf die Personalauswahl bei der Entwicklung der Instrumente. Diese Versuche sind im konkreten Fall möglicherweise nur begrenzt erfolgreich, z. B. weil sich im konkreten Prozess durch die Interaktion der einzelnen Rahmenbedingungen Konstellationen ergeben, die dazu führen, dass die Intentionen der bildungspolitischen Ebene sich nicht durchsetzen.

9.2.3 Implikationen der jeweiligen Rahmenbedingungen

Im Folgenden werden die Ergebnisse dieser Arbeit in Bezug auf mögliche Ursachen und Folgen ausführlich diskutiert. Dabei wird erneut auf die drei Dimensionen zurückgegriffen, die nach Spillane et al. (2002) die Implementation von politischen Vorgaben prägen. Im Anschluss werden Bezüge zu den theoretischen Überlegungen aus der Bewertungssoziologie hergestellt.

„Persönliche Attribute“ In Bezug auf die Dimension „persönliche Attribute“ wird das Potenzial des Sensemaking-Ansatzes deutlich, weil dieser im Gegensatz zu institutionellen Ansätzen den kognitiven Deutungsprozess des Individuums sowie Attribute des Individuums in den Blick nimmt (vgl. ebd.). Im Folgenden wird auf zwei Aspekte näher eingegangen: die Relevanz der eigenen Praxiserfahrung sowie die geringe Thematisierung von Emotionen.

a) Relevanz der eigenen Praxiserfahrung erschwert Innovationen

Wie bereits oben dargestellt zeigte sich bei denjenigen Instrumenteentwickler_innen⁶, die auch als Inspektor_innen tätig sind oder als Lehrkraft tätig waren bzw. sind, dass sie bei der Deutung von Inklusion bzw. der Entwicklung der Instrumente auf ihre eigene Praxiserfahrung als Lehrkraft oder auf Beobachtungen, die sie im Rahmen ihrer Inspektionstätigkeit gemacht haben, zurückgriffen. Dies bestätigt die Annahmen von Spillane et al. (ebd., S. 393), wonach Sensemaker bei Deutungsprozessen bestehende Wissensbestände und Erfahrungen nutzen (siehe auch DeMatthews 2012). Des Weiteren bestätigt dies auch Arbeiten, die sich damit auseinandersetzen, welche Relevanz biographische Erfahrungen für den professionellen Bereich haben (vgl. Fabel und Tiefel 2004).

Hierbei ist zu bedenken, dass neue Vorgaben nach Spillane et al. meistens so gedeutet werden, dass sie zu bestehenden Wissensbeständen und Überzeugungen kohärent sind und somit ein Paradigmenwechsel im Denken selten bis gar nicht stattfindet (vgl. Spillane et al. 2002, S. 396): „We are told old habits die hard. Mental scripts are equally resilient“ (Spillane 2004, S. 78).

⁶ Andere Interviewpartner hatten eine Leitungsrolle inne bzw. waren wissenschaftliche Referent_innen des jeweiligen Instituts.

Dieses Phänomen zeigt sich auch bei den Instrumenteentwickler_innen, da die Entwicklung von Instrumenten zu Inklusion zwar teilweise als Herausforderung wahrgenommen wurde, aber von ihnen wurde nicht thematisiert, dass dies dazu geführt hätte, dass bestehende Wissensbestände und Überzeugungen kritisch hinterfragt oder revidiert worden wären. Des Weiteren wurde in den Interviews auch nicht angesprochen, dass die Entwicklung von Instrumenten zu Inklusion dazu geführt hätte, dass bereits vorhandene Messinstrumente hinterfragt oder geändert worden wären.

Für die Einführung von politischen Vorgaben, die große Veränderungen bzw. ein Umdenken in Bezug auf bestehende Praktiken implizieren, ist es demnach notwendig, dass die Instrumenteentwickler_innen vorab eine konkrete Idee erhalten, welche Implikationen dies für die Praxis hat, beispielsweise durch Hospitationen an Best-practice-Schulen im In- und Ausland. Indem diese Erfahrungen gemacht werden und bei der Deutung von Inklusion darauf zurückgegriffen wird, könnte dies zur Folge haben, dass andere Wissensbestände, auf die üblicherweise bei der Deutung von Inklusion zurückgegriffen wird, z. B. Erfahrungen in Sonderschulen, an Relevanz verlieren.

Eine weitere Möglichkeit wäre auch, dass bei der Kommunikation von politischen Vorgaben an Wertvorstellungen und Wissensbestände angeknüpft wird, die bereits vorhanden sind, aber andere Deutungsprozesse initiieren. Dies wäre der Fall, wenn Inklusion nicht als Integration von Schüler_innen mit Behinderung von Seiten der Kultusministerien kommuniziert wird, sondern als Erweiterung des Konzepts „Umgang mit Heterogenität“ oder als „Schutz von Menschenrechten für alle“ gesehen wird.

In Bezug auf die Entwicklung von Messinstrumenten stellt sich die Frage, ob bei der Entwicklung der anderen Messinstrumente, die von der Schulinspektion zur Messung von Schulqualität genutzt werden, die eigenen Erfahrungen und Beobachtungen auch solch eine große Rolle gespielt haben. Dies wurde im Rahmen der Arbeit empirisch nicht untersucht, jedoch ist zu bedenken, dass die Instrumenteentwickler_innen im Fall von Inklusion – nach eigenen Angaben in den Interviews – kaum auf andere Quellen, z. B. wissenschaftliche Literatur oder konkrete Vorgaben von bildungspolitischer Seite, zurückgreifen konnten. Es liegt nahe, dass dies bei den anderen Instrumenten nicht der Fall war und demnach könnte es sein, dass die eigenen Erfahrungen eine weniger wichtige Rolle eingenommen haben. Dennoch ist es wahrscheinlich, dass auch bei der Entwicklung der anderen Messinstrumente tief verankerte Überzeugungen zu Schulqualität, die auf vorherigen Erfahrungen basieren, nicht automatisch durch externen Input in Frage gestellt oder revidiert wurden, sondern den Deutungsprozess geprägt haben.

b) Emotionale Ebene wird kaum thematisiert außer in Bezug auf Schulqualität

Spillane et al. führen darüber hinaus aus, dass Gefühle und Wertüberzeugungen bei der Implementation von politischen Vorgaben prägend wirken (vgl. Spillane et al. 2002). Wie bereits im Methodenkapitel ausführlich diskutiert zeigte sich, dass dieser Aspekt in den Interviews kaum thematisiert wurde. Hierbei lässt sich nicht abschließend beantworten, ob die geringe Thematisierung impliziert, dass jene Annahme von Spillane et al. nicht bestätigt werden kann. Alternativ ist es auch möglich, dass die geringe Thematisierung sich aus dem empirischen Design der Arbeit ergibt, z. B. geringe Thematisierung im Leitfaden sowie Ansprache der Instrumenteentwickler_innen in ihrer professionellen Rolle (nähere Ausführung → Kapitel 6). Gleichwohl wird an diesem Beispiel eine Grenze des Sensemaking-Ansatzes deutlich: So lässt sich nur begrenzt sicherstellen, ob die Sensemaker sich aller Rahmenbedingungen, die beim Deutungsprozess für sie relevant waren, bewusst sind und diese auch im Rahmen der Interviews genannt haben.

„**Situativer Kontext**“ Neben der Dimension „persönliche Attribute“ ist noch der „situative Kontext“ zu nennen, der nach Spillane et al (2002) die Deutung von politischen Vorgaben prägt. In diesem Zusammenhang wird im Folgenden auf die strukturelle Verortung von Schulinspektionen sowie den Mangel an demokratischer Legitimation näher eingegangen.

a) Strukturelle Verortung von Schulinspektionen im Bildungssystem legt das Handlungsmuster „Streben nach Akzeptanz“ nahe

Ausgehend von dem oben beschriebenen Handlungsmuster „Streben nach Akzeptanz“ lassen sich Bezüge zu anderen Studien, die mit dem Sensemaking-Ansatz oder dem Educational-Governance-Ansatz arbeiten, herstellen. Beiden Ansätzen ist als Prämisse gemein, dass formal bestehende Machtverhältnisse nur begrenzt erklären können, wie es zu bestimmten Entscheidungen und Handlungen kommt. So zeigt sich bei der Entwicklung der Instrumente, dass schulische Vertreter – obwohl sie in der formalen Hierarchie im Vergleich zu Schulinspektionen niedriger anzusiedeln sind – bei der Entwicklung der Instrumente in vielen Fällen beteiligt waren bzw. dass ihre Interessen bei der Formulierung von Anforderungen antizipiert wurden. Von Seiten der Instrumenteentwickler_innen wurde argumentiert, dass die Instrumente bei schulischen Vertretern Akzeptanz erfahren müssten, damit die Impulsfunktion der Schulinspektion Wirkung zeigen könnte. Die Orientierung an schulischen Vertretern ist nicht überraschend, da viele Instrumenteentwickler_innen früher selbst als Lehrkräfte tätig waren bzw. dies zum Teil weiterhin in Teilzeit sind. Somit liegt es nahe, dass sie sich auf persönlicher Ebene den schulischen Vertretern verbunden fühlen. Jener Aspekt zeigt sich auch in der empirischen Erhebung von Sowada und Dederling, wonach ein Inspektor argumentiert: *„Also, ich sage mal: Als ehemalige Schulleiter waren wir ja nicht daran interessiert, irgendwie jetzt Angst und Schrecken zu verbreiten“* (Sowada und Dederling 2016, S. 190). Demnach wird hier die Annahme von Spillane et al. (2002, S. 409) bestätigt, dass der berufliche Hintergrund der Sensemaker – in diesem Fall die Tätigkeit als Lehrkraft oder Schulleitung – das Sensemaking prägt.

Des Weiteren wird in einzelnen Interviews thematisiert, dass die Akzeptanz von Seiten der schulischen Vertreter von Relevanz sei, weil dies sich wiederum darauf auswirken würde, wie sich die Kultusministerien zu den Schulinspektionen positionieren. In diesem Zusammenhang ist auch zu erwähnen, dass von Seiten des Kultusministeriums und der schulischen Vertreter häufig ähnliche Interessen verfolgt wurden: So wurden Schulinspektionen zwar von Seiten der Ministerien eingeführt, um die Qualität an Einzelschulen zu überprüfen, aber in den Interviews wurde thematisiert, dass die Ministerien sich der Gefahr bewusst seien, dass zu hohe Anforderungen an Schule zugleich dazu führen könnten, dass die Ministerien aufgrund schlechter Evaluationsergebnisse bei den Schulen selbst unter Rechtfertigungsdruck geraten (siehe auch Piezunka 2019).

In diesem Zusammenhang stellt sich die Frage, was es bedeutet, wenn schulischen Vertretern, z. B. Schulleitungen oder Lobbyverbänden von Lehrkräften, bei der Entwicklung von Instrumenten so viel Mitspracherecht gegeben wird. So hat dies zur Folge, dass die Gruppe von Personen, die evaluiert wird, (mit-)entscheiden kann, welche Anforderungen an sie formuliert werden. Hier bietet sich ein Verweis auf die Kritik von Böttcher et al. zu Selbstevaluation an: *„Schließlich muss gesichert sein, dass die Schule auch dort evaluiert wird, wo es <weh tut>“* (Böttcher et al. 2006, S. 15). Im Fall der Schulinspektion ist es zwar nicht die Einzelschule, der die konkrete Evaluation überlassen wird, aber die drohende Gefahr ist ähnlich: Indem schulische Vertreter bei der Entwicklung der Instrumente aktiv miteinbezogen werden, kann dies zur Folge haben, dass die Anforderungen den Schulen „nicht weh tun“ und somit auch nicht dazu führen, dass Schulentwicklungsprozesse initiiert werden. So wurde in Bezug auf

die Instrumente zu Inklusion in den Interviews explizit der Fall beschrieben, dass die Anforderungen aus Rücksichtnahme auf die Schulen niedriger formuliert wurden. Es fehlt bislang an empirischen Untersuchungen, inwiefern dieser Aspekt auch bei den anderen Messinstrumenten eine Rolle gespielt hat.

Das Streben nach Akzeptanz gegenüber dem Kultusministerium, schulischen Vertretern und der Schulaufsicht lässt sich möglicherweise durch die strukturelle Verankerung von Schulinspektionen erklären. So befinden sich die Schulinspektionen in einem Abhängigkeitsverhältnis zum Kultusministerium, da dieses über deren Existenz bzw. Fortführung entscheidet und hierfür berücksichtigt, welche Rückmeldungen es von Seiten der schulischen Vertreter erhält. Daran anknüpfend stellt sich jedoch die Frage, inwiefern „Streben nach Akzeptanz“ als dominantes Handlungsmuster möglicherweise dazu führt, dass Schulinspektionen ihre Existenzberechtigung im Bildungssystem verlieren (nähere Ausführung Piezunka 2019). So kann der hohe Grad an Mitspracherechten durch schulische Vertreter zwar dazu führen, dass die Schulinspektion von Seiten der schulischen Vertreter akzeptiert und auf Einzelschulebene Entwicklungsprozesse initiiert werden. Dies hätte zur Folge, dass die Impulsfunktion von Schulinspektionen erfüllt wird. Gleichwohl kann die Kontrollfunktion, die auch durch Schulinspektionen realisiert werden soll, nur begrenzt realisiert werden, wenn schulischen Vertretern Mitspracherecht gegeben wird. In diesem Zusammenhang ist zu bedenken, dass die Kontrollfunktion ein Alleinstellungsmerkmal von Schulinspektionen darstellt, da sie in der Bildungsadministration – neben der Schulaufsicht – die Einzigen⁷ sind, die (mithilfe von empirischen Datenerhebungen) die Qualität von Einzelschulen messen (nähere Ausführung → Kapitel 3).

Die Impulsfunktion kann dagegen auch durch andere Maßnahmen, z. B. interne Evaluationen, realisiert werden. Die stärkere Orientierung an der Impulsfunktion und der Relevanzverlust der Kontrollfunktion, kann dazu führen, dass die Existenzberechtigung von Schulinspektionen nicht mehr ausreichend legitimiert ist. Diese These liegt insofern nahe, als beispielsweise die Abschaffung der Schulinspektionen in Hessen und in Rheinland-Pfalz dazu geführt hat, dass mehr Ressourcen in den Bereich der internen Evaluation gesteckt wurden.

b) Mangel an demokratischer Legitimation aufgrund von Nichteinbeziehung bestimmter Akteure

Bislang wurde thematisiert, welche Rahmenbedingungen die Entwicklung von Instrumenten zu Inklusion geprägt haben. Es ist auch zu berücksichtigen, welche Personengruppen oder Rahmenbedingungen nicht prägend wirkten. So zeigte sich bei der Analyse, dass in den Interviews nur einmal berichtet wurde, dass der Kontakt zur Behindertenbeauftragten gesucht wurde. Auch andere Lobbyverbände, wie Elternvereine, wurden als beteiligte Akteure in den Interviews nicht genannt. Dies widerspricht dem aus der internationalen Behindertenbewegung stammenden Motto „Nothing about us without us“ (vgl. Degener 2009, S. 202). Des Weiteren zeigte sich, dass nicht nur Repräsentanten von Minderheiten, sondern generell Eltern und Schüler_innen bei der Entwicklung der Instrumente nicht involviert waren.

Hierbei lässt sich ein Kritikpunkt, den Rüb und Straßheim gegenüber evidenzbasierter Politik und ihrem Verweis auf wissenschaftliche Evidenzen formulieren, auf das Handeln der Instrumenteentwickler_innen von Schulinspektionen übertragen: Der „Kampf um die Anerkennung von bestimmten normativen Prinzipien und der damit verbundenen Politisierung der (medialen) Öffentlichkeit, [tritt, ap] in den Hintergrund (...)“ (Rüb und Straßheim 2012, S. 378). Hierbei merkt Straßheim an:

⁷ Auch die Schulaufsicht nimmt im Bildungssystem eine Kontrollfunktion gegenüber Schulen wahr, aber sie erhebt keine empirischen Daten, um die Qualität von Schulen zu überprüfen.

„This tendency poses problems for democracies in terms of both <epistemic equality> in the sense of equal rights for everyone to be taken seriously when making knowledge claims and <epistemic neutrality> in the sense of impartiality of the state in ensuring open and fair discussion in the face of rival claims“ (Straßheim 2015, S. 321; siehe auch Biesta 2009, S. 39).

Die fehlende Einbeziehung der oben genannten Gruppen ist aus demokratietheoretischer Perspektive demnach kritisch zu sehen: So wird dadurch verhindert, dass unterschiedliche Stimmen angehört werden können und eine transparente Debatte darüber, was gute Schule ist, entstehen kann.

„Politische Vorgabe“ Als weitere Dimension ist die politische Vorgabe zu nennen bzw. deren Repräsentation, z. B. durch Begleitmaterial wie Empfehlungen der Kultusministerien oder der UN. Wie bereits oben erwähnt zeigte sich in den Interviews, dass sich die Instrumenteentwickler_innen kaum an den konkreten Inhalten von Art. 24 BRK orientierten. Des Weiteren wurden auch andere gesetzliche Vorgaben z. B. auf Bundeslandebene oder konkrete Maßnahmenpakete von politischer Seite in den Interviews kaum thematisiert.

In diesem Zusammenhang stellt sich die Frage, inwiefern dies ein Widerspruch dazu ist, dass die politische Vorgabe als solche – im Rahmen dieser Arbeit „Inklusion“ – bei Spillane et al. ein eigenes Element darstellt (vgl. Spillane et al. 2002, S. 393). Spillane et al. argumentieren selbst, dass die politische Vorgabe auch dem Element „situativer Kontext“ zugeordnet werden könnte. In dem Fall würde der Deutungsprozess von den persönlichen Attributen des Sensemakers, z. B. Wertüberzeugungen, und dem situativen Kontext geprägt werden. Für die Benennung als eigenes Element spricht jedoch, dass diese im Vergleich zu anderen Rahmenbedingungen in Bezug auf den Deutungsprozess von besonderer Relevanz ist. Dies lässt sich in Bezug auf den hier betrachteten Fall insofern bestätigen, als die gesetzlichen Vorgaben dazu geführt haben, dass bei den Schulinspektionen Prozesse in Richtung Inklusion initiiert wurden. Des Weiteren machen sie deutlich, dass durch dieses Element primär in den Blick genommen wird, wie eine politische Vorgabe nach außen kommuniziert und von Seiten der Instrumenteentwickler_innen interpretiert wird (vgl. ebd., S. 416).

Darüber hinaus zeigt sich an diesem Beispiel eine Herausforderung des Sensemaking-Ansatzes: So liefern die von Spillane et al. (2002) aufgeführten Elemente (persönliche Attribute, situativer Kontext sowie politische Vorgabe) beim empirischen Design, d. h. bei der Entwicklung des Leitfadens sowie der Auswertung, als Kategoriensystem wichtige inhaltliche Hinweise. Gleichwohl zeigt sich während der Auswertung auch, dass die einzelnen Rahmenbedingungen nur begrenzt analytisch getrennt und einzelnen Elementen zugeordnet werden können. Dies lässt sich u.a. damit erklären, dass die einzelnen Rahmenbedingungen im tatsächlichen Deutungsprozess der einzelnen Sensemaker miteinander interagieren.

a) Politische Vorgaben als „one sentence statement“

Politische Vorgaben werden nach Spillane et al. häufig als „one sentence statement cast as goals or objectives“ (ebd., S. 416) kommuniziert. In Bezug auf Inklusion als politische Vorgabe liegt es nahe, dass Art. 24 BRK im Sinne eines „one sentence statement“ von vielen Instrumenteentwickler_innen (aber nicht allen) folgendermaßen gelesen wurde: „Schüler_innen mit Behinderung haben ein Recht darauf, eine Regelschule zu besuchen“. So zeigt die empirische Analyse (→ Abschnitt 7.2), dass viele Schulinspektionen Inklusion mit der Differenzlinie „Behinderung“ verknüpfen – sei es in der Deutung von Inklusion oder bei der Auswahl an Schulen, die zu Inklusion evaluiert werden (Ausnahme: Schleswig-Holstein und Hamburg). Dies steht

im Kontrast zu den Setzungen, die im Rahmen dieser Arbeit vorgenommen wurden, wonach Inklusion die Teilhabe und Nichtdiskriminierung aller Schüler_innen anstrebt.

b) Erklärungen für die fehlende Auseinandersetzung mit rechtlichen Vorgaben

Die fehlende Auseinandersetzung mit den Anforderungen, die sich aus den politischen Vorgaben ergeben, lässt sich möglicherweise dadurch erklären, dass es von Seiten der Bildungspolitik an Beschreibungen fehlt, welche konkreten Implikationen die jeweiligen Vorgaben für die einzelnen Bereiche, z. B. Unterrichtsebene oder Interaktion zwischen Lehrkräften und Schüler_innen, haben. In diesem Zusammenhang weisen Spillane et al. (2002) darauf hin, dass dichte Beschreibungen, die als Ergänzung zu den konkreten Gesetzesvorgaben veröffentlicht werden, einen Beitrag leisten können, damit politische Vorgaben so umgesetzt werden, wie sie von politischer Seite intendiert sind.

Wenn von höherer bildungspolitischer Ebene diese Aufgabe der Konkretisierung nicht übernommen wird, hat dies zur Folge, dass die Schulinspektionen stärker als Policymaker fungieren können: Indem sie Inklusion als politische Vorgabe in den Messinstrumenten operationalisieren, konkretisieren sie, was die Implementation von Inklusion für die Einzelschule bedeutet.

c) Bundesländerübergreifende Harmonisierung der Instrumente

Ausgehend von dem Argument, dass es für alle Bundesländer gilt, dass sie sich durch die Ratifizierung von verschiedenen UN-Konventionen zur Umsetzung von Inklusion verpflichtet haben, würde eine Harmonisierung der Instrumente naheliegen. Demnach würden alle Schulinspektionen die gleichen Anforderungen an Inklusion formulieren. In diesem Fall müssten sich jedoch alle beteiligten Bundesländer auf einen gemeinsamen Nenner einigen. Dies könnte bei der Operationalisierung von Inklusion zur Folge haben, dass möglichst niedrigschwellige Anforderungen formuliert werden, damit diese bundesländerübergreifend konsensfähig sind. Dies wird für die Implementation von Inklusion als politische Vorgabe von der Autorin dieser Arbeit insofern kritisch gesehen, als dass es - basierend auf den Ergebnissen dieser Arbeit - unter denjenigen bleiben, die sich aus den UN-Konventionen und bundesländerspezifischen Schulgesetzen ergeben. Gleichwohl führt das föderale System in Deutschland dazu, dass sich die Bundesländer bei der Umsetzung von Inklusion unterscheiden und somit unterschiedliche Anforderungen an die Schulen formuliert werden, was im Widerspruch dazu steht, dass es sich um eine international verabschiedete Vorgabe handelt.

9.2.4 Bewertungssoziologische Überlegungen

Abschließend wird nun näher betrachtet, welche Schlussfolgerungen sich aus der empirischen Analyse generell für die Entwicklung von Instrumenten zur Messung von Qualität ziehen lassen.

Legitimation des eigenen Handelns In der Analyse der Interviews sowie in den Dokumenten zeigte sich das Phänomen, dass bei der allgemeinen Beschreibung der Entstehung der Messinstrumente häufig darauf verwiesen wurde, dass diese auf einem allgemeingültigen Verständnis von guter Schule basieren (nähere Ausführung → Exkurs in 7.5) bzw. auf Erkenntnisse aus der Bildungsforschung zurückgegriffen wird. Jene Aussagen stehen im Kontrast zu den Ergebnissen, die bei der Betrachtung eines konkreten Falls, der Entwicklung und der Inhalte der Instrumente zu Inklusion, herausgearbeitet wurden. Beispielsweise wird in den Interviews thematisiert, dass die bestehenden Beiträge aus der Bildungsforschung bei der Entwicklung von Instrumenten zu Inklusion kaum für hilfreich erachtet wurden und die Analyse der Instrumente

machte deutlich, dass sich die Bundesländer darin unterscheiden, wie sie die Umsetzung von Inklusion messen.

Im Folgenden wird nun mit Bezügen zu bewertungssoziologischen Arbeiten (nähere Ausführung → Kapitel 4) diskutiert, wie sich die Abweichung zwischen konkreten Handlungen und den Begründungen des Handelns erklären lassen (doings vs. sayings). So verweist Desrosières (2009) auf das Phänomen „double consciousness“ von Favret-Saada. Danach ist es manchen Funktionsrollen inhärent, dass sie zwei Haltungen verkörpern müssen, die nicht miteinander kompatibel sind. Angewandt auf die Arbeiten von Statistiker_innen bedeutet dies:

„In their everyday practice, statisticians are plunged into a world of conventions, which they record or shape themselves. The fact that the measurement results from this sequencing of conventional decisions is therefore self-evident to them. Later, however they change that without realizing it, and speak a realist language when addressing the outside world. There is no ambiguity here: in both cases, statisticians have done their job conscientiously. (...) <I'm well aware that my statistics result from conventions, yet I believe in a reality that I am asked to quantify“ (ebd.: 321- 322).

Demnach ist Statistiker_innen – wie bereits dargestellt – bei der Entwicklung von Indikatoren bewusst, dass diese nicht naturgegeben und objektiv sind, sondern sie hierfür auf bestimmte soziale Konventionen zurückgreifen. Es wäre demnach auch möglich, andere Indikatoren für die Abbildung von Wirklichkeit zu verwenden. Zugleich müssen sie, wenn sie sich auf bestimmte Indikatoren festgelegt haben, der Außenwelt vermitteln, dass nur die ausgewählten Indikatoren es ermöglichen, die Realität abzubilden. So ist dies eine notwendige Grundbedingung, damit die jeweiligen Indikatoren als legitimes Messinstrument verstanden und als Grundlage für weiteres Handeln genutzt werden.

In diesem Zusammenhang lassen sich Anknüpfungspunkte zum Handeln der Instrumenteentwickler_innen in Schulinspektionen herstellen: Auch diese sind sich – wie oben dargestellt – dessen bewusst, dass bei der Entwicklung von Instrumenten eine Vielzahl von unterschiedlichen Rahmenbedingungen prägend wirkt und dass es unterschiedliche Formen gibt, wie beispielsweise die Umsetzung von Inklusion gemessen werden könnte. Dennoch begründen sie ihre Instrumente häufig damit, dass sie auf einem allgemein anerkannten Verständnis von Schulqualität basieren würden. Es liegt nahe, dass diese Argumentation – wie bei den von Desrosières beschriebenen Statistiker_innen – gewählt wird, um die eigenen Instrumente zu legitimieren. Sie entgehen somit der Herausforderung, dass die Instrumente von anderen Akteuren in Frage gestellt werden können: *„Auch in unserem Verfahren wollen wir nicht anfangen mit der Schule über die Instrumente oder die herausgekommenen Indikatoren zu diskutieren, sondern wir haben es im Qualitätsrahmen festgelegt, worum es gehen soll“* (Int. 18_12). Demnach könnte man jene Aussagen der Instrumenteentwickler_innen nach Thévenot als Forminvestition verstehen, d. h. als Bemühung darum, dass die eigenen Instrumente über eine bestimmte Reichweite sowie eine bestimmte Dauer an Gültigkeit gewinnen (vgl. Thévenot 2009).

Neben dem Verweis auf ein allgemeingültiges Verständnis von Schulqualität werden die Instrumente häufig damit legitimiert, dass sie sich an den jeweiligen Qualitätsrahmen der Bundesländer orientieren (Int. 6_3; 5_14; 15_12; 9_7; 5_13); *„Also es is ja im Grunde genommen so, wir sind durch den Orientierungsrahmen komplett legitimiert in unserem Handeln. Das is ja schonmal der Vorteil. Wir können immer sagen: So, Schulinspektion is dafür da den Orientierungsrahmen abzubilden und zu erheben wie weit die Schulen mit Blick auf die dort formulierten Qualitätsansprüche schon sind“* (Int. 13_16).

Des Weiteren wird – wie bereits oben dargestellt – auf die Bildungsforschung verwiesen. Jener Verweis lässt sich nach Straßheim damit erklären, dass diese von bildungspolitischer Seite

im Sinne von evidence-basierter Politik vermeintlich für apolitisch bzw. objektiv gehalten wird (vgl. (Straßheim 2015, S. 321). Damit geht der Eindruck einher, dass Forscher_innen eine Form von „reiner Wahrheit“ beanspruchen könnten. Es entsteht der Eindruck einer „scheinbare[n] Norm- und Interessenlosigkeit von Evidenzen“ (Rüb und Straßheim 2012, S. 378) und die „Umstrittenheit und auch das Auch-Anders-Sein-Können politischer Urteile und Legitimitäten“ (ebd., S. 378) geht verloren. Angewendet auf Schulinspektionen impliziert die Legitimation der Instrumente durch die Verweise auf die Bildungsforschung einen ähnlichen Effekt wie die Behauptung, dass diese auf einem allgemeingültigen Verständnis von Schulqualität basieren: Den Instrumenten wird Objektivität zugeschrieben und diese werden für alternativlos angesehen.

Wenn im politischen Bereich auf „wissenschaftliche Expertise“ verwiesen wird, sind es meistens die Ergebnisse von quantitativen Analysen bzw. deren Visualisierung, die dazu dienen das eigene Handeln zu rechtfertigen. Rüb und Straßheim (ebd., S. 378) beschreiben dies als „kalkulativ-axiomatische Evidenz“. Im Fall von Schulinspektionen geht es jedoch nicht um bestehende Daten aus der empirischen Bildungsforschung, stattdessen stehen die Instrumenteentwickler_innen vor der Herausforderung, ihre eigene Datenerhebung, d. h. die Auswahl der Messinstrumente, zu begründen. Dabei verweisen sie auf bestehende Indikatorensets bzw. konkrete Bildungsforscher wie Hattie, Helmke und Meyer. Auch dieser Verweis auf Forschung fällt nach Rüb und Straßheim unter den Evidenzbegriff: So beschreiben sie dies als „soziale Evidenz“. Demnach wird einzelnen Akteuren – in diesem Fall Wissenschaftler_innen – aufgrund von Erfahrungen und Detailwissen Expertenstatus zugeschrieben (vgl. ebd., S. 384) und durch den Verweis auf jene „Expert_innen“ wird das eigene Handeln legitimiert. Straßheim warnt jedoch vor dieser Gleichsetzung von Wissenschaft und Expertise: „Experts tend to neglect evidence which does not come from their profession, are skeptical of novel views, and fail to see the schemes of values that is part of their judgement“ (Straßheim 2015, S. 320). Demnach sind die Perspektiven von Bildungsforscher_innen bzw. ihre Wissensbestände begrenzt und auch sie forschen nicht wertfrei bzw. „objektiv“. Dies ist insbesondere nicht der Fall, wenn es um so hoch normative Themen wie Schulqualität und Inklusion geht.

In diesem Zusammenhang ist auch zu berücksichtigen, welche Wissenschaftler_innen von Seiten der Schulinspektionen als Expert_innen herangezogen wurden. So zeigte sich in vielen Fällen, dass unter anderem bereits vorhandene Kontakte, geographische Nähe oder Bekanntheitsgrad ausschlaggebend dafür waren, wer eingeladen wurde.

Die Bemühung um vermeintliche Objektivität ist stimmig mit dem Bild der Schulinspektionen als ein Akteur, der evidenzbasiert Rückmeldungen gibt und sich somit von anderen Akteuren, die subjektive Einschätzungen treffen, z. B. Eltern, abgrenzen möchte (→ Kapitel 3). Um sich von subjektiven Überzeugungen abgrenzen zu können und an Autorität zu gewinnen, kann daher der Verweis auf ein allgemein gültiges Verständnis von Schulqualität und auf die Bildungsforschung dazu dienen, die eigenen Instrumente zu legitimieren. Dabei liegt es nahe, dass die Instrumenteentwickler_innen tatsächlich an sich selbst den Anspruch formulieren, die Ergebnisse der Bildungsforschung als Grundlage zu nutzen und sich an einem allgemein gültigen Verständnis von guter Schule zu orientieren. Indem sie jedoch im Entwicklungsprozess verschiedenen Akteuren und Rahmenbedingungen gerecht werden müssen und es generell keinen Konsens darüber gibt, was unter einer guten Schule zu verstehen ist, scheitern sie an diesem Anspruch.

In diesem Zusammenhang lässt sich auch ein Bezug zu den Rechtfertigungsordnungen von Boltanski und Thévenot (2007) herstellen. Indem das eigene Vorgehen bei der Entwicklung

von Instrumenten damit gerechtfertigt wird, dass auf bestehende Indikatorensets bzw. auf einzelne Bildungsforscher_innen verwiesen wird, zeigen sich Anknüpfungspunkte zur „Welt der Reputation“ sowie zur „Welt der Industrie“ (vgl. Diaz-Bone 2015, S. 144). Demnach verweisen die Instrumenteentwickler_innen auf die Arbeiten von Bildungsforscher_innen, weil sie davon ausgehen, dass diese auch von der Öffentlichkeit als Expert_innen in Bezug auf Instrumente anerkannt werden. Anhand der systematischen Sichtung der Literatur und der Orientierung an bestehenden Indikatorensets wird die Orientierung an der „Welt der Industrie“ sichtbar.

Hierbei zeigt sich zugleich, auf welche Rechtfertigungsmuster nicht verwiesen wird: So wird in den Interviews kaum auf die „staatsbürgerliche Welt“ verwiesen, um das eigene Handeln zu legitimieren. Diese kennzeichnet sich durch Werte wie Partizipation und Anerkennung von unterschiedlichen Interessen (vgl. ebd., S. 153). Jenes Rechtfertigungsmuster würde sichtbar werden, wenn die Instrumenteentwickler_innen ihr eigenes Handeln damit begründen, dass sie versuchen die Interessen von verschiedenen Gruppen, wie Eltern, Schüler_innen sowie Lobbyverbänden, miteinzubeziehen.

Fehlendes Wissen über die Entwicklung von Instrumenten zur Messung von Qualität

Bislang gibt es sowohl im Bildungsbereich als auch in anderen Bereichen, z. B. Medizin, Ingenieurwesen, usw., nur vereinzelt Untersuchungen, die sich damit auseinandersetzen, wie Instrumente zur Messung von Qualität entstehen und welche Rahmenbedingungen hierbei prägend wirken. Demnach lassen sich kaum Bezüge zu bestehenden Untersuchungen bzw. theoretischen Überlegungen herstellen (nähere Ausführung → Kapitel 4). Es gibt jedoch insofern Anknüpfungspunkte zu den empirischen Arbeiten von Lamont (2010) und Chenhall et al. (2013), als auch hier die persönlichen Attribute derjenigen Personen, die die Instrumente entwickeln, eine große Rolle spielen. Darüber hinaus zeigen sich Überschneidungen darin, dass die Mitarbeiter_innen der Schulinspektionen bei der Entwicklung der Instrumente antizipieren, welche Wirkung diese erzielen könnten bzw. sollten.

Zusammenfassend: Die Anwendung des Sensemaking-Ansatzes für die Analyse der relevanten Rahmenbedingungen bei der Entwicklung von Messinstrumenten ermöglicht es, dass dadurch verschiedene Aspekte in den Blick genommen werden konnten, z. B. bestehende Strukturen und Abhängigkeiten, die sich aus dem situativen Kontext ergeben, sowie Attribute auf individueller Ebene.

In Bezug auf die theoretische Ebene zeigte sich retrospektiv, dass viele der Komponenten, die Heid benennt, um Qualitätsurteile analysieren zu können, auch dabei helfen zu verstehen, welche Rahmenbedingungen bei der Entwicklung von Messinstrumenten relevant sind. Hierbei ist zu berücksichtigen, dass die Messinstrumente als solche bei Heid auch eine eigene Komponente darstellen (vgl. Heid 2013, S. 411): die sogenannten „Kriterien“. Als weitere Komponenten nennt Heid unter anderem das „Subjekt der Qualitätszuschreibung“. Diese spiegeln sich in der Dimension „persönliche Attribute der Instrumenteentwickler_innen“ im Sensemaking-Ansatz wider. Des Weiteren ist der „Zweck der Qualitätsbeurteilung“ (vgl. ebd., S. 411) zu nennen, der sich unter anderem in den Funktionen der Schulinspektionen ausdrückt, die beim Sensemaking-Ansatz dem situativen Kontext zugeschrieben werden.

Auch die „Adressaten des Qualitätsurteils“, die im hier untersuchten Fall die Einzelschule bzw. die schulischen Vertreter sind, prägen die Rahmenbedingungen bei der Entwicklung der Instrumente. Darüber hinaus nennt Heid als Komponente den „Gegenstand der Qualitätsbeurteilung“ – in diesem Falle die Schule. Diese spielt bei der Entwicklung der Instrumente eine Rolle, wenn es darum geht, welche gesellschaftliche Funktionen Schule zu erfüllen hat und

inwiefern diese messbar sind. Als letzte Komponente nennt Heid den Kontext bzw. die Randbedingungen, in denen ein Qualitätsurteil (vgl. ebd., S. 411) getroffen wird. Hier lassen sich unter anderem Bezüge zur strukturellen Verankerung von Schulinspektionen im Bildungssystem und den sich daraus ergebenden Abhängigkeiten herstellen. Demnach bieten sich die Komponenten von Heid als Analysemodell an, um Hinweise darauf zu erhalten, welche Rahmenbedingungen von Relevanz sind. Die Verknüpfung mit dem Sensemaking-Ansatz ist jedoch notwendig, um den Deutungsprozess und die Entwicklung der Instrumente nachvollziehen zu können. So trifft dieser Annahmen darüber, inwiefern die Rahmenbedingungen miteinander interagieren und welche Tendenzen es bei kognitiven Deutungsprozessen gibt, z. B. Festhalten an bestehenden Überzeugungen.

10 Fazit

Im Rahmen der Arbeit wurde die Entwicklung von Messinstrumenten am Beispiel Inklusion näher untersucht. Dabei leistete diese Arbeit nicht nur einen Beitrag zu der Frage, wie Messinstrumente entstehen und inwiefern Inklusion als Verständnis von guter Schule in den Instrumenten berücksichtigt wird. Die Arbeit ermöglicht auch Aussagen zur Implementation von Inklusion als politische Vorgabe auf Bundesländerebene in den einzelnen Bildungsadministrationen. Im Folgenden wird dargestellt, welchen Beitrag die Arbeit in Hinblick auf diese Themenbereiche für Praxis und Forschung darstellt.

In Bezug auf die Forschung lässt sich Folgendes festhalten: Die Arbeit bestätigt empirische Untersuchungen zur Umsetzung von Inklusion, indem sie nachweist, dass es auch auf der Ebene der Bildungsadministration keinen Konsens darüber gibt, wie Inklusion definiert wird. Die Arbeit geht noch einen Schritt weiter, indem sie mithilfe des Sensemaking-Ansatzes analysiert, welche Rahmenbedingungen die Deutungen von Inklusion prägen. Des Weiteren knüpft die Arbeit an bewertungssoziologische Studien an: Bislang gibt es kaum Forschung zu der Frage, welche Rahmenbedingungen die Entwicklung von Instrumenten zur Messung von Qualität prägen. Hier wurde – nach dem Wissen der Autorin – erstmalig der Sensemaking-Ansatz auf den Gegenstand der Messinstrumente angewandt. Dadurch konnten verschiedene Rahmenbedingungen, wie persönliche Attribute, aber auch der situative Kontext, in den Blick genommen werden. Die Betrachtung davon, wie Messinstrumente entstehen, lässt wiederum Aussagen darüber zu, warum sich bestimmte Wertvorstellungen durchgesetzt haben und andere nicht. In diesem Fall wurde deutlich, welche Relevanz den persönlichen Erfahrungen der Instrumenteentwickler_innen und der strukturellen Verankerung der Schulinspektion zuzuschreiben ist. Hierbei stellt sich die Frage, inwiefern dies auf andere Messinstrumenten – sei es im Bildungsbereich oder in anderen Bereichen – übertragen werden kann oder inwiefern die Schulinspektionen in Deutschland einen Sonderfall darstellen.

In Bezug auf weitere Forschung würde es sich demnach anbieten, zu untersuchen, inwiefern sich bei der Entwicklung von Messinstrumenten in anderen Fällen ähnliche Handlungsmuster wie die „Relevanz der eigenen Erfahrung“ oder „Streben nach Akzeptanz“ sichtbar werden. So würde es naheliegen, die Entwicklung von Instrumenten zu untersuchen, die als Grundlage für weitreichende Entscheidungen dienen, z. B. wenn das Statistische Bundesamt neue Indikatoren entwickelt oder wenn Indikatoren im Dienstleistungsbereich entstehen, die der Bewertung von Produkten dienen. Aus methodischer Perspektive würde es sich anbieten, hier nicht nur auf Interviews, sondern auch auf teilnehmende Beobachtungen zurückzugreifen (siehe auch DeMatthews 2012). Dadurch könnte herausgearbeitet werden, auf welche Rechtfertigungsordnungen im Laufe des Prozesses zurückgegriffen wird. Darüber hinaus könnten kollektive Aushandlungsprozesse auf Bundesländerebene stärker in den Blick genommen werden. So lag im Rahmen dieser Arbeit der Fokus primär auf den individuellen Sensemakern und ihrer Wahrnehmung von kollektiven Aushandlungsprozessen. Um noch stärker die Unterschiede zwischen den bundesländerspezifischen Schulinspektionen herausarbeiten zu können, wäre es notwendig gewesen, in den einzelnen Bundesländern mit mehr Personen zu sprechen und bei relevanten Sitzungen dabei zu sein. In Bezug auf die Messinstrumente wäre es auch interessant, deren Implementation in der Praxis näher zu betrachten, beispielsweise wie die Instrumente zu Inklusion bei der Inspektion gedeutet werden, wie sie im Inspektionsbericht festgehalten werden und wie diese von Seiten der Schule rezipiert werden.

Als weiteres Forschungsfeld bietet sich die Analyse von Algorithmen an. So werden im Alltag immer häufiger Entscheidungen getroffen, die auf computerbasierten Algorithmen basieren, z. B. Personalauswahl in großen Unternehmen. Dabei zeigt sich beispielsweise, dass jene Algorithmen das Potenzial haben, soziale Ungleichheiten zu reproduzieren. Beispielsweise beschreibt O'Neill (2018), dass der Rückgriff auf Algorithmen dazu führe, dass die Versicherungsbeiträge von People of Color aufgrund der niedrigeren Lebenserwartung höher angesetzt wurden als die Beiträge von Weißen. Die Ursachen für die geringe Lebenserwartung wurden jedoch bei der algorithmusbasierten Berechnung nicht berücksichtigt. Die Analyse von Algorithmen bietet sich insofern an, als diese – ähnlich wie quantitative Daten – häufig als objektiv angesehen werden bzw. nach außen als objektiv präsentiert werden (~mathwashing) (vgl. ebd.). Demnach liegt es nahe, verstärkt zu untersuchen, wie sich Algorithmen auf Frage der Teilhabe und Diskriminierung auswirken können und was mögliche Ursachen hierfür sind.

Aus **bildungspolitischer Perspektive** ist von Relevanz, dass die Umsetzung von Inklusion in den einzelnen Bundesländern unterschiedlich gedeutet bzw. gemessen wird. So unterscheiden sich die Bundesländer darin, welche Anforderungen sie an Schulen formulieren bzw. inwiefern das Erfüllen der Anforderungen für die Schulen verpflichtend ist. Dabei wird bei der Bildungsadministration auch eine Unsicherheit darüber sichtbar, wie die BRK zu interpretieren ist. Dies wird beispielsweise darin deutlich, dass es zwischen den Bundesländern variiert, welche Gruppe in Bezug auf Inklusion genannt wird bzw. wer unter den Schutz der BRK fällt (vgl. Piezunka et al. 2018). In diesem Zusammenhang ist auch in Betracht zu ziehen, dass Deutungsprozesse – in diesem Fall zu Inklusion – nach dem Sensemaking-Ansatz nie abgeschlossen sind und weiterhin stattfinden. Demnach ist es möglich, dass zukünftige Instrumente der Schulinspektionen bzw. Instrumente, die nach dem Erhebungszeitpunkt entwickelt wurden, adaptierte Deutungen von Inklusion repräsentieren.

In Bezug auf Schulinspektionen lässt sich darüber hinaus festhalten, dass diese mit ihren gewählten Formen der Datenerhebung (Fokus auf die Prozessebene) sowie den gewählten Inhalten eine Alternative zu anderen empirischen Messungen von Qualität im Bildungsbereich darstellen. So zeigt sich insbesondere im angloamerikanischen Bereich die Relevanz von Lernstandserhebungen, die auch in Deutschland in Form des IQB-Ländervergleichs und der PISA-Studie eine große Rolle spielen. Diese konzentrieren sich auf die Outputebene und nehmen überwiegend die Qualifizierungsfunktion von Schule in den Blick. Ausgehend von der Frage, welche Werteordnungen sich bei der Entwicklung von Instrumenten durchsetzen, lässt sich festhalten, dass Inklusion als Teilaspekt von Schulqualität in den Instrumenten der Schulinspektionen berücksichtigt wird, aber dass sich die Bundesländer darin unterscheiden, wie sie Inklusion interpretieren, welche Relevanz sie dem Thema beimessen und wie sie dies messen.

Des Weiteren konnte mithilfe des Sensemaking-Ansatzes gezeigt werden, wie es zu den unterschiedlichen Verständnissen von Inklusion gekommen ist und wie die einzelnen Rahmenbedingungen miteinander interagiert haben. Dies ist für die Praxis insofern von Relevanz, als sich daraus beispielsweise für die ministeriale Ebene zum einen die Erkenntnis ergibt, dass sie durch die Auswahl der Instrumenteentwickler_innen oder die Einbeziehung von bestimmten Gruppen beeinflussen lässt, wie die Umsetzung von Inklusion von Seiten der Inspektionen gedeutet und gemessen wird. Zum anderen ergibt sich aus der Arbeit auch die Erkenntnis, dass jene Prozesse nur begrenzt von außen beeinflusst werden können, weil eine Vielzahl von Rahmenbedingungen bzw. deren Interaktion eine Rolle spielen. So können auch ähnliche Rahmenbedingungen aufgrund von unterschiedlichen Gewichtungen oder Machtkonstellationen zu unterschiedlichen Deutungen von Inklusion führen.

Darüber hinaus zeigte sich im Rahmen der Arbeit die Relevanz der strukturellen Verankerung von Schulinspektionen im Bildungssystem: Beispielsweise legt diese ein „Streben nach Akzeptanz“ nahe. In Anbetracht der Tatsache, dass viele Schulinspektionen in den letzten Jahren abgeschafft wurden bzw. aktuell pausieren, stellt sich die Frage, ob die bestehenden Abhängigkeiten und Interessen, die sich aus der strukturellen Verankerung ergeben, auch eine Ursache dafür sind, dass die Existenzberechtigung von Schulinspektionen in den letzten Jahren häufig in Frage gestellt wurde. In diesem Zusammenhang ist auch zu diskutieren, inwiefern Schulinspektionen primär eine Kontroll- oder eine Impulsfunktion ausüben sollten bzw. ob beide Funktionen miteinander kompatibel sind.

Abschließend zeigt sich durch die Analyse der Instrumente zu Inklusion, dass es im Bundesländervergleich bei Schulinspektionen keinen Konsens darüber gibt, was unter einer guten Schule verstanden wird und wie dies gemessen werden soll. Es ist jedoch wichtig, in der Bildungspraxis sowie in der Forschung eine transparente Debatte darüber zu führen und hierfür unterschiedliche Personen bzw. Interessengruppen einzubeziehen:

„But if we are not explicit about our views about the aims and ends of education – if we do not tackle the questions as to what constitutes good education head on – we run the risk that statistics and league tables will make these decisions for us. We therefore need to keep the question of purpose – the question of what constitutes good education – central in our educational discussions and wider endeavours“ (Biesta, 2009, S. 44).

Literatur

- Ahyoud, Nasiha, Kwesi Aikins, Joshua, Bartsch, Samera, Bechert, Naomi, Gyamerah, Daniel und Wagner, Lucienne (2018). *Wer nicht gezählt wird, zählt nicht: Antidiskriminierungs- und Gleichstellungsdaten in der Einwanderungsgesellschaft – eine anwendungsorientierte Einführung*. URL: www.vielfaltentscheidet.de/publikationen.
- Aikins, Joshua Kwesi, Bartsch, Samera, Gyamerah, Daniel und Wagner, Lucienne (2018). *Diversität in öffentlichen Einrichtungen: Antidiskriminierungs- und Gleichstellungsdaten in der Praxis - Ergebnisse einer Piloterhebung unter Führungskräften der Berliner Verwaltung und landeseigenen Unternehmen*. Hrsg. von Vielfalt entscheidet – Diversity in Leadership. (Besucht am 06. 04. 2018).
- Ainscow, Mel (2005). "Developing inclusive education systems: what are the levers for change?" In: *Journal of Educational Change* 6.2, S. 109–124.
- Ainscow, Mel, Booth, Tony und Dyson, Alan (2006). *Improving Schools, Developing Inclusion*. Abingdon: Routledge.
- Alicke, Tina und Linz-Dinchel, Kathrin (2012). *Inklusive Gesellschaft - Teilhabe in Deutschland*. Hrsg. von Institut für Sozialarbeit und Sozialpädagogik e.V.
- Aljets, Enno (2015). *Der Aufstieg der empirischen Bildungsforschung: Ein Beitrag zur institutionalistischen Wissenschaftssoziologie*. Wiesbaden: VS Verlag für Sozialwissenschaften.
- Altrichter, Herbert, Thomas Brüsemeister und Jochen Wissing, Hrsg. (2007). *Educational Governance: Handlungskoordination und Steuerung im Bildungssystem*.
- Altrichter, Herbert und Kemethofer, David (2016). "Stichwort: Schulinspektoren". In: *Zeitschrift für Erziehungswissenschaft* 19.3, S. 487–508.
- Alvesson, Mats (2003). "Beyond Neopositivists, Romantics, and Localists: A reflexive approach to interviews in organizational research". In: *Academy of Management Review* 28.1, S. 13–33.
- Amrhein, Bettina (2011). *Inklusion in der Sekundarstufe: Eine empirische Analyse*. Bad Heilbrunn: Verlag Julius Klinkhardt.
- Anagnostopoulos, Dorothea, Lingard, Bob und Sellar, Sam (2016). "Argumentation in Educational Policy Disputes: Competing Visions of Quality and Equity". In: *Theory Into Practice* 55.4, S. 342–351.
- Appelt, Doreen (2016). *Sensemaking und Sensegiving in der Sanierung: Zur Einflussnahme auf die Bedeutungskonstruktion durch das Management*. Wiesbaden: VS Verlag für Sozialwissenschaften.
- Arbeitsgruppe Schulinspektion (2016). *Schulinspektion als Steuerungsimpuls? Ergebnisse aus Forschungsprojekten*. Bd. 25. Educational Governance. Wiesbaden: VS Verlag für Sozialwissenschaften.
- Artelt, Cordula (2007). "Externe Evaluation und einzelschulische Entwicklung - ein zukunftsreiches Entwicklungsverständnis". In: *Qualität von Schule*. Hrsg. von Jürgen van Buer und Cornelia Wagner. Frankfurt am Main: Peter Lang, S. 131–140.
- Autorengruppe Bildungsberichterstattung (2016). *Bildung in Deutschland 2016. Ein indikatorengestützter Bericht mit einer Analyse zu Bildung und Migration*.
- Avenarius, Hermann (2006). *Regelungen zur Rolle der Schulaufsicht und zum Umfang ihrer Aufgaben und Befugnisse*. Visselhövede-Jeddingen.
- (2012). "Auf dem Weg zur inklusiven Schule? Auszüge eines viel beachteten Vortrags von Prof. Dr. Hermann Avenarius anlässlich des Ganztagsschulkongresses im November 2011". In: *Schulverwaltung / Nordrhein-Westfalen* 2.3, S. 66–70.
- Bacher, Johann, Winklhofer, Ursula und Teubner, Markus (2007). "Partizipation von Kindern in der Grundschule". In: *Kinderleben - Start in die Grundschule*. Hrsg. von Christian Alt. Schriften des Deutschen Jugendinstituts. Wiesbaden: VS Verlag für Sozialwissenschaften, S. 271–299.
- Balzer, Nicole (2014). *Spuren der Anerkennung: Studien zu einer sozial- und erziehungswissenschaftlichen Kategorie*. Dordrecht: Springer.
- Balzer, Nicole und Ricken, Norbert (2010). "Anerkennung als pädagogisches Problem: Markierungen im erziehungswissenschaftlichen Diskurs". In: *Anerkennung* 1, S. 35–87.
- Bastian, Johannes (2009). "Schülerbeteiligung lernen: Lern- und Schulkultur gemeinsam entwickeln". In: *Pädagogik* 7.8, S. 5–9.
- Bayerisches Staatsministerium für Unterricht und Kultus (2005). *Externe Evaluation an Bayerns Schulen: Das Konzept, die Instrumente, die Umsetzung*. Hrsg. von Qualitätsagentur am Staatsinstitut für Schulqualität und Bildungsforschung.
- (2010). *Externe Evaluation an Bayerns Schulen*. URL: https://www.isb.bayern.de/schulartuebergreifendes/qualitaetsicherung-schulentwicklung/evaluation/externe_evaluation/broschuere/ (besucht am 05. 11. 2015).
- Becker, Howard Saul (2006). *Tricks of the trade: How to think about your research while you're doing it*. Chicago guides to writing, editing, and publishing. Chicago: The University of Chicago Press.

- Beckert, Jens und Musselin, Christin (2013). "Introduction". In: *Constructing Quality: The Classification of Goods in Markets*. Hrsg. von Jens Beckert und Christin Musselin. Oxford University Press, S. 1–30.
- Behnke, Kristin und Steins, Gisela (2017). "Principals' reactions to feedback received by school inspection: A longitudinal study". In: *Journal of Educational Change* 18.1, S. 77–106.
- Behörde für Schule und Berufsbildung Hamburg (2012). *Orientierungsrahmen Schulqualität und Leitfaden*. (Besucht am 04. 02. 2014).
- Bellmann, Johannes, Duzevic, Doris, Schweizer, Sebastian und Thiel, Corrie (2016). "Nebenfolgen neuer Steuerung und die Rekonstruktion ihrer Genese". In: *Zeitschrift für Pädagogik* 62.3, S. 381–401.
- Bellmann, Johannes und Müller, Thomas (2011). "Evidenzbasierte Pädagogik - ein Déjà-vu? Einleitende Bemerkungen zur Kritik eines Paradigmas". In: *Wissen was wirkt*. Hrsg. von Johannes Bellmann und Thomas Müller. Wiesbaden: VS Verlag für Sozialwissenschaften, S. 9–32.
- Bellmann, Johannes, Schweizer, Sebastian und Thiel, Corrie (2016). "Nebenfolgen Neuer Steuerung unter Bedingungen von <low-stakes> und <no-stakes>: Qualitative und quantitative Befunde einer Untersuchung in vier Bundesländern". In: *Steuerung im Bildungssystem*, S. 208–237.
- Berliner Senatsverwaltung für Bildung, Jugend und Wissenschaft (2012). *Zweite Runde Schuleinspektion in Berlin: Handbuch ab Schuljahr 2011/12*. (Besucht am 07. 10. 2013).
- Bernhard, Nadine (2014). "Durch Europäisierung zu mehr Durchlässigkeit? Veränderungsdynamiken des Verhältnisses von beruflicher Bildung zur Hochschulbildung in Deutschland und Frankreich". Diss. Berlin: Freie Universität Berlin.
- Bertelsmann-Stiftung (2014). *Update Inklusion - Datenreport zu den aktuellen Entwicklungen*. Hrsg. von Bertelsmann-Stiftung.
- Bezem, Pablo und Jacobsen, Rebecca (2018). *School inspector decision making in three countries: Argentina, Germany, and The Netherlands*. Mexico City.
- Bielefeldt, Heiner (2009). *Zum Innovationspotenzial der UN-Behindertenrechtskonvention*. Hrsg. von Deutsches Institut für Menschenrechte. Berlin.
- Biesta, Gert (2009). "Good education in an age of measurement: On the need to reconnect with the question of purpose in education". In: *Educational Assessment, Evaluation and Accountability* 21.1, S. 33–46.
- (2010). "Valuing What We Measure or Measuring What We Value? On the Need to Engage with the Question of Purpose in Educational Evaluation, Assessment, and Measurement". In: *Evaluation, Bildung und Gesellschaft*. Hrsg. von Wolfgang Böttcher, Jan Nikolas Dicke und Nina Hogrebe. Münster, New York, NY, München, Berlin: Waxmann, S. 35–46.
- (2011). "Warum What works nicht funktioniert: Evidenzbasierte pädagogische Praxis und das Demokratiedefizit der Bildungsforschung". In: *Wissen was wirkt*. Hrsg. von Johannes Bellmann und Thomas Müller. Wiesbaden: VS Verlag für Sozialwissenschaften, S. 95–121.
- (2015). "Education, Measurement and the Professions: Reclaiming a space for democratic professionalism in education". In: *Educational Philosophy and Theory* 49.4, S. 315–330.
- Biewer, Gottfried (2009). *Grundlagen der Heilpädagogik und inklusiven Pädagogik*. Bad Heilbrunn: Verlag Julius Klinkhardt.
- (2012). "Die neue Welt der Bildungsstandards und ihre erziehungswissenschaftliche Rezeption aus der Perspektive einer inklusiven Pädagogik". In: *Vierteljahresschrift für Heilpädagogik und ihre Nachbargebiete* 81, S. 9–21.
- Bittner, Melanie (2011). *Geschlechterkonstruktionen und die Darstellung von Lesben, Schwulen, Bisexuellen, Trans* und Inter* (LSBTI) in Schulbüchern: Eine gleichstellungsorientierte Analyse von Melanie Bittner*. Hrsg. von Max-Träger-Stiftung. (Besucht am 17. 02. 2014).
- Black-Hawkins, Kristine, Florian, Lani und Rouse, Martyn (2007). *Achievement and inclusion in schools*. London, New York: Routledge.
- Blanck, Jonna Milena, Edelstein, Benjamin und Powell, Justin J. W. (2013). "Persistente schulische Segregation oder Wandel zur inklusiven Bildung? Die Bedeutung der UN-Behindertenrechtskonvention für Reformprozesse in den deutschen Bundesländern". In: *Swiss Journal of Sociology* 39.2, S. 267–292.
- BLLV (2017). *Schulleitungen klagen an: Mehr geht nicht: Bllv schreibt Brandbrief an Ministerpräsident*. URL: <https://www.bllv.de/vollstaendiger-artikel/news/schulleitungen-klagen-an-mehr-geht-nicht-mehr/> (besucht am 04. 12. 2017).
- Bloem, Simone (2016). *Die PISA-Strategie der OECD: Zur Bildungspolitik eines globalen Akteurs*. Neue Politische Ökonomie der Bildung. Weinheim: Beltz Juventa.
- Boban, Ines und Hinz, Andreas (2003). *Index für Inklusion. Lernen und Teilhabe in der Schule der Vielfalt entwickeln*. URL: <http://www.eenet.org.uk/resources/docs/Index%20German.pdf> (besucht am 16. 04. 2012).
- (2011). "Inklusion - mehr als ein neuer pädagogischer Begriff. Eine internationale Schulreformenentwicklung". In: *Schulmagazin* 5 - 10 79.12, S. 7–10.

- Boban, Ines, Hinz, Andreas, Plate, Elisabeth und Tiedeken, Peter (2014). "Inklusion in Worte fassen - eine Sprache ohne Kategorisierungen?" In: *Inklusion und Chancengleichheit*. Hrsg. von Saskia Schuppener. Bad Heilbrunn: Verlag Julius Klinkhardt, S. 19–24.
- Bogner, Alexander, Littig, Beate und Menz, Wolfgang (2014). *Interviews mit Experten: Eine praxisorientierte Einführung*. Wiesbaden: VS Verlag für Sozialwissenschaften. (Besucht am 03. 04. 2017).
- Böhm-Kasper, Oliver (2006). "Schulische und politische Partizipation von Jugendlichen. Welchen Einfluss haben Schule, Familie und Gleichaltrige auf die politische Teilhabe Heranwachsender?" In: *Diskurs Kindheits- und Jugendforschung* 1.3, S. 353–368.
- Böhm-Kasper, Oliver und Selders, Odette (2013). "Schulinspektionen sollten regelmäßig durchgeführt werden? Ländervergleichende Analyse der Wahrnehmung und Akzeptanz von Schulinspektionsverfahren". In: *Die deutsche Schule - DDS. Beiheft* 12: Evidenzbasierte Steuerung im Bildungssystem? : Befunde aus dem BMBF-SteBis-Verbund.
- Boltanski, Luc und Thévenot, Laurent (1999). "The Sociology of Critical Capacity". In: *European Journal of Social Theory* 2.3, S. 359–377.
- (2007). *Über die Rechtfertigung: Eine Soziologie der kritischen Urteilskraft*. Hamburg: Hamburger Edition.
- Booth, Tony und Ainscow, Mel (2002). *Index for Inclusion: developing learning and participation in schools*. Hrsg. von Centree for Studies on Inclusive Education. (Besucht am 12. 03. 2017).
- Bos, Wilfried, Dederig, Kathrin, Holtappels, Heinz-Günter, Müller, Sabine und Rösner, Ernst (2007). "Schulinspektion in Deutschland - eine kritische Bestandsaufnahme". In: *Qualität von Schule*. Hrsg. von Jürgen van Buer und Cornelia Wagner. Frankfurt am Main: Peter Lang, S. 241–259.
- Böttcher, Wolfgang (2009). "Was leisten Evaluationen für die Qualitätsentwicklung?" In: *Lernen aus Evaluationsergebnissen*. Hrsg. von Thorsten Bohl und Hanna Kiper. Bad Heilbrunn: Verlag Julius Klinkhardt, S. 207–217. URL: http://www.gew-nds.de/Aktuell/archiv_august_09/Boettcher_Qualitaetsentwicklung.pdf (besucht am 10. 08. 2016).
- Böttcher, Wolfgang, Holtappels, Heinz Günter und Brohm, Michaela (2006). "Evaluation im Bildungswesen". In: *Evaluation im Bildungswesen*. Hrsg. von Wolfgang Böttcher, Heinz Günter Holtappels und Michaela Brohm. Grundlagentexte Pädagogik. Weinheim: Juventa Verlag, S. 7–21.
- Böttcher, Wolfgang und Keune, Miriam (2010). "Funktionen und Effekte der Schulinspektion: Ausgewählte nationale und internationale Forschungsbefunde". In: *Evaluation, Bildung und Gesellschaft*. Hrsg. von Wolfgang Böttcher, Jan Nikolas Dicke und Nina Hogrebe. Münster, New York, NY, München, Berlin: Waxmann, S. 151–164.
- (2011). "Wirkungsanalyse der Schulinspektion: Empirische Ergebnisse aus Deutschland". In: *Wirkungen und Wirksamkeit der externen Schulevaluation*. Hrsg. von Carsten Quesel, Norbert Landwehr, Peter Steiner und Vera Husfeldt. Bern: hep der Bildungsverlag, S. 123–144.
- Böttger-Beer, Manuela und Koch, Erik (2008). "Externe Schulevaluation in Sachsen - ein Dialog zwischen Wissenschaft und Praxis". In: *Bildungsmonitoring und Bildungscontrolling in nationaler und internationaler Perspektive*. Hrsg. von Wolfgang Böttcher, Wilfried Bos, Hans Döbert und Heinz Günter Holtappels. Münster u. a.: Waxmann, S. 253–264.
- Botz, Lieselotta, Heese, Sarah und Schirow, Olga (2011). "Qualitätskriterien bei Schulinspektionen". In: *Schulinspektion*. Hrsg. von Matthias von Saldern. Schule in Deutschland. Norderstedt: Books on Demand, S. 113–133.
- Bowker, Geoffrey C. und Star, Susan Leigh (2000). *Sorting Things out: Classifications and its consequences*. Cambridge, Massachusetts: The MIT Press.
- Bradbury, Alice und Roberts-Holmes, Guy (2017). "Creating an Ofsted story: The role of early years assessment data in schools' narratives of progress". In: *British Journal of Sociology of Education* 38.7, S. 943–955.
- Breidenstein, Georg (2017). "Das Theorem der 'Selektionsfunktion der Schule'". In: *Leistung als Paradigma*. Hrsg. von Sabine Reh und Norbert Ricken. Wiesbaden: VS Verlag für Sozialwissenschaften, S. 307–327.
- Brüggemann, Nele, Jäkel, Susan und Riemer, Tabea (2011). "Schulinspektion - ein historischer Vergleich". In: *Schulinspektion*. Hrsg. von Matthias von Saldern. Schule in Deutschland. Norderstedt: Books on Demand, S. 32–63.
- Brüsemeister, Thomas (2007). "Steuerungsakteure und ihre Handlungslogiken im Mehrebenensystem der Schule". In: *Governance, Schule und Politik*. Hrsg. von Jürgen Kussau und Thomas Brüsemeister. Wiesbaden: VS Verlag für Sozialwissenschaften, S. 63–95.
- Bundesverwaltungsamt (2010). *Deutsche Schulen weltweit unterwegs zur Exzellenz: Bund-Länder-Inspektion Bilanz 2010*. (Besucht am 11. 01. 2018).
- (2014). *Bund-Länder-Inspektion 2.0: BLI-Folgezyklus*.
- Burzan, Nicole und Ronald Hitzler, Hrsg. (2017). *Theoretische Einsichten: Im Kontext empirischer Arbeit*. Wiesbaden: VS Verlag für Sozialwissenschaften.
- Campbell, Carol (2002). "Conceptualizations and definitions of inclusive schooling". In: *Developing inclusive schooling*. Hrsg. von Carol Campbell. Bedford Way papers. London: Institute of Education, University of London, S. 11–34.

- Castel, Patrick (2009). "What's Behind a Guideline? Authority, Competition and Collaboration in the French Oncology Sector". In: *Social Studies of Science* 39.5, S. 743–764.
- Chenhall, Robert, Hall, Matthew und Smith, David (2013). "Performance measurement, modes of evaluation and the development of compromising accounts". In: *Accounting, Organizations and Society* 38, S. 268–287. (Besucht am 22. 04. 2016).
- Coburn, Cynthia (2001). "Collective Sensemaking about Reading: How Teachers mediate Reading policy in their professional communities". In: *Educational Evaluation and Policy Analysis* 23.2, S. 145–170.
- (2006). "Framing the Problem of Reading Instruction: Using Frame Analysis to Uncover the Microprocesses of Policy Implementation". In: *American Educational Research Journal* 43.3, S. 343–379.
- (2016). "What's Policy Got to Do with It? How the Structure-Agency Debate Can Illuminate Policy Implementation". In: 122.3.
- Colyvas, Jeannette A. (2012). "Performance Metrics as Formal Structures and through the Lens of Social Mechanisms: When Do They Work and How Do They Influence?" In: *American Journal of Education* 118.2, S. 167–197.
- Crenshaw, Kimberle (1989). *Demarginalizing the Intersection of Race and Sex: A Black Feminist Critique of Antidiscrimination Doctrine, Feminist Theory, and Antiracist Politics*. URL: http://www-polisci.tamu.edu/upload_images/4/Crenshaw-Demarginalizing.pdf (besucht am 16. 07. 2012).
- Dederig, Kathrin (2012). "Schulinspektion als wirksamer Weg der Systemsteuerung?" In: *Zeitschrift für Pädagogik* 58, S. 69–88. (Besucht am 14. 05. 2014).
- Dederig, Kathrin und Müller, Sabine (2011). "School improvement through inspections? First empirical insights from Germany". In: *Journal of Educational Change* 12.3, S. 301–322.
- Degener, Theresia (2009). "Die UN-Behindertenrechtskonvention als Inklusionsmotor". In: *Recht der Jugend und des Bildungswesens* 52.2, S. 200–219.
- Degener, Theresia und Mogge-Grotjahn, Hildegard (2012). „All inclusive“? Annäherungen an ein interdisziplinäres Verständnis von Inklusion". In: *Soziale Inklusion*. Hrsg. von Hans-Jürgen Balz, Benjamin Benz und Carola Kuhlmann. Wiesbaden: VS Verlag für Sozialwissenschaften, S. 59–77.
- Degenhardt, Sven (2008). „Gute Schule“ in der Bundesrepublik Deutschland - auch eine „Gute Schule“ für blinde und sehbehinderte Schülerinnen und Schüler? URL: <https://www.ew.uni-hamburg.de/de/ueber-die-fakultaet/personen/degenhardt/files/2008-sd-guteschule-citydruck-pdf.pdf> (besucht am 18. 10. 2014).
- DeMatthews, David (2012). "Principal Sensemaking of Inclusion: A Multi-case study of five urban School Principals". Dissertation. Maryland: University of Maryland.
- Desrosières, Alain (2000). "Measurement and Its Uses: Harmonization and Quality in Social Statistics". In: *International Statistical Review* 68.2, S. 173–187.
- (2005). *Die Politik der großen Zahlen: Eine Geschichte der statistischen Denkweise*. Berlin Heidelberg: VS Verlag für Sozialwissenschaften.
- (2009). "How to be Real and Conventional: A Discussion of the Quality Criteria of Official Statistics". In: *Minerva* 47.3, S. 307–322.
- Deutsche Gesellschaft für Demokratiepädagogik e.V. (2014). *ABC der Demokratiepädagogik*.
- Deutsche UNESCO-Kommission (2014). *Inklusion: Leitlinien für die Bildungspolitik*. URL: http://www.unesco.de/fileadmin/medien/Dokumente/Bildung/2014_Leitlinien_inklusive_Bildung.pdf (besucht am 11. 11. 2014).
- Deutsches Institut für Menschenrechte (2011). *Stellungnahme der Monitoring-Stelle: Eckpunkte zur Verwirklichung eines inklusiven Bildungssystems (Primarstufe und Sekundarstufen I und II)*. URL: http://www.institut-fuer-menschenrechte.de/uploads/tx_commerce/stellungnahme_der_monitoring_stelle_eckpunkte_z_verwirklichung_eines_inklusive_bildungssystems_31_03_2011.pdf (besucht am 07. 07. 2014).
- (2014). *Inklusive Bildung: Schulgesetze auf dem Prüfstand. Vorabfassung der Studie*. URL: http://www.institut-fuer-menschenrechte.de/uploads/tx_commerce/Vorabfassung_Studie_Inklusive_Bildung_Schulgesetze_auf_dem_Pruferstand.pdf (besucht am 20. 03. 2014).
- (2015). *Parallelbericht an den UN-Fachausschuss für die Rechte von Menschen mit Behinderungen*.
- Deutsches Kinderhilfswerk e.V. (2013). *Umfrage zum politischen Engagement von Jugendlichen*.
- Dewey, John (1939). *Theory of valuation*. Chicago Ill: The University of Chicago Press.
- Diaz-Bone, Rainer (2015). *Die Economie des conventions: Grundlagen und Entwicklungen der neuen französischen Wirtschaftssoziologie*. Wiesbaden: VS Verlag für Sozialwissenschaften.
- Die Antidiskriminierungsstelle des Bundes (2013). *Diskriminierung im vorschulischen und schulischen Bereich: Eine sozial- und erziehungswissenschaftliche Bestandsaufnahme*.
- Die Beauftragte der Bundesregierung für Migration, Flüchtlinge und Integration (2015). *Schulbuchstudie Migration und Integration*. Hrsg. von Georg-Eckert-Institut und Stiftung Universität Hildesheim.
- Diedrich, Martina (2008). "Demokratische Schulkultur. Messung und Effekte". Diss. URL: http://www.pedocs.de/volltexte/2010/2670/pdf/Diss_MD_Manuskript_2008_08_18_D_A.pdf (besucht am 17. 03. 2016).

- Diedrich, Martina und Pietsch, Marcus (2013). "Wie unterstützt eine Schulinspektion die Entwicklung schulischer Qualität?: Aufbereitung von Daten zur Ergebnismeldung in der Praxis". In: *Journal für Schulentwicklung* 17.1, S. 19–23.
- Diefenbach, Heike (2010). "Bildungschancen und Bildungs(miss)erfolg von ausländischen Schülern oder Schülern aus Migrantenfamilien im System schulischer Bildung". In: *Bildung als Privileg*. Hrsg. von Rolf Becker und Wolfgang Lauterbach. Wiesbaden: VS Verlag für Sozialwissenschaften, S. 217–241.
- Diehl, Claudia, Christian Hunkler und Cornelia Kristen, Hrsg. (2016). *Ethische Ungleichheiten im Bildungsverlauf: Mechanismen, Befunde, Debatten*. Wiesbaden: VS Verlag für Sozialwissenschaften.
- Dietrich, Fabian (2016). "Schulinspektion als Steuerungsimpuls zur Schulentwicklung? Objektiv-hermeneutische Governance-Analysen zur Handlungskoordination im Kontext der Schulinspektionen". In: *Schulinspektion als Steuerungsimpuls?* Hrsg. von Arbeitsgruppe Schulinspektion. Educational Governance. Wiesbaden: VS Verlag für Sozialwissenschaften, S. 119–168.
- (2018). "Konturen einer Rekonstruktiven Governanceforschung: Zu einer rekonstruktiven Perspektivierung von Schule als Mehrebenensystem und deren Stellenwert im Kontext der Schul- und Professionalisierungsforschung". In: *Rekonstruktive Bildungsforschung*. Hrsg. von Martin Heinrich und Andreas Wernet. Wiesbaden: VS Verlag für Sozialwissenschaften, S. 73–94.
- Dietrich, Fabian und Lambrecht, Maïke (2012). "Menschen arbeiten mit Menschen: Schulinspektion und die Hoffnung auf den zwanglosen Zwang der <besseren Evidenz>". In: *Die deutsche Schule* 104.3, S. 57–70.
- Disability News (2017). *Ofsted report 'shows government is ignoring alarm bells on inclusive education'*. URL: <https://www.disabilitynewsservice.com/ofsted-report-shows-government-is-ignoring-alarm-bells-on-inclusive-education/> (besucht am 27. 12. 2017).
- Ditton, Hartmut (2000). "Qualitätskontrolle und Qualitätssicherung in Schule und Unterricht. Ein Überblick zum Stand der empirischen Forschung". In: *Zeitschrift für Pädagogik* 41, S. 73–92. (Besucht am 21. 01. 2014).
- Ditton, Hartmut und Müller, Andreas (2011). "Schulqualität". In: *Empirische Bildungsforschung*. Hrsg. von Heinz Reinders, Hartmut Ditton, Cornelia Gräsel und Burkhard Gniewosz. Lehrbuch. Wiesbaden: VS Verlag für Sozialwissenschaften, S. 99–111.
- Döbert, Hans (2009). "Indikatorenentwicklung im Rahmen der Bildungsberichterstattung in Deutschland". In: *Indikatorenentwicklung für den nationalen Bildungsbericht „Bildung in Deutschland“*, S. 9–22.
- Döbert, Hans, Rürup, Matthias und Dederig, Kathrin (2008). "Externe Evaluation von Schulen in Deutschland - die Konzepte der Bundesländer, ihre Gemeinsamkeiten und Unterschiede". In: *Externe Evaluation von Schulen*. Hrsg. von Hans Döbert und Kathrin Dederig. Münster, New York, München und Berlin: Waxmann, S. 63–151.
- Döring, Nicola und Bortz, Jürgen (2016). *Forschungsmethoden und Evaluation in den Sozial- und Humanwissenschaften*. Heidelberg: Springer.
- Durkheim, Emile und Mauss, Marcel (1987). "Über einige primitive Formen von Klassifikation: Ein Beitrag zur Forschung der kollektiven Vorstellungen". In: *Schriften zur Soziologie der Erkenntnis*. Hrsg. von Emile Durkheim, Hans Joas und Michael Bischoff. Frankfurt am Main: Suhrkamp, S. 169–256.
- Ehren, Melanie (2016). "Introducing School Inspections". In: *Methods and Modalities of Effective School Inspections*. Hrsg. von Melanie Ehren. Wiesbaden: Springer, S. 1–18.
- Ehren, Melanie und Scheerens, Jaap (2015). "Evidenzbasierte Referenzrahmen zur Schulqualität als Grundlage von Schulinspektionen". In: *Schulinspektion in Hamburg*. Hrsg. von Marcus Pietsch, Barbara Scholand und Klaudia Schulte. Münster, New York, NY, München, Berlin: Waxmann, S. 233–272.
- Ehren, Melanie und Visscher, Adrie J. (2006). "Towards a Theory on the Impact of School Inspections". In: *British Journal of Educational Studies* 54.1, S. 51–72.
- Eikel, Angelika (2006). *Demokratische Partizipation in der Schule*. URL: http://degede.de/fileadmin/public/partizipation_sfoerderung/01_Demokr._Partizipation_in_der_Schule.pdf (besucht am 04. 10. 2015).
- Emcke, Carolin (2010). *Kollektive Identitäten: Sozialphilosophische Grundlagen: Dissertation*. Frankfurt am Main, New York: Campus-Verlag.
- Espeland, Wendy Nelson (2015). "Narrating Numbers". In: *The world of indicators*. Hrsg. von Richard Rottenburg. Cambridge studies in law and society. Cambridge: Cambridge University Press, S. 56–75.
- Espeland, Wendy Nelson und Sauder, Michael (2007). "Rankings and Reactivity: How Public Measures Recreate Social Worlds". In: *American Journal of Sociology* 113.1, S. 1–40.
- (2016). *Engines of anxiety: Academic rankings, reputation, and accountability*. New York: Russell Sage Foundation.
- Espeland, Wendy Nelson und Stevens, Mitchell L. (1998). "Commensuration as a Social Process". In: *Annual Review of Sociology* 24.1, S. 313–343.
- European Agency for Development in Special Needs Education (2011). *Mapping the Implementation of Policy for Inclusive Education: An Exploration of Challenges and Opportunities for Developing Indicators*. URL: http://www.european-agency.org/sites/default/files/mipie-report_MIPIE-Report.pdf (besucht am 10. 10. 2013).

- Fabel, Melanie und Tiefel, Sandra (2004). "Biographie als Schlüsselkategorie qualitativer Professionsforschung - eine Einleitung". In: *Biographische Risiken und neue professionelle Herausforderungen*. Hrsg. von Melanie Fabel und Sandra Tiefel. VS Verlag für Sozialwissenschaften, S. 11–40.
- Fallarchiv Universität Kassel (2018). *INTAKT*. URL: <http://www.fallarchiv.uni-kassel.de/projektdateien-intakt/> (besucht am 28. 06. 2018).
- Felder, Franziska (2012). *Inklusion und Gerechtigkeit: Das Recht behinderter Menschen auf Teilhabe*. Bd. 956. Campus Forschung. Frankfurt am Main, New York: Campus-Verlag.
- Fend, Helmut (1994). "Was ist eine gute Schule?" In: *Was ist eine gute Schule?* Hrsg. von Klaus-Jürgen Tillmann. PB-Bücher. Hamburg: Bergmann + Helbig, S. 14–25.
- (2000). "Qualität und Qualitätssicherung im Bildungswesen. Wohlfahrtsstaatliche Modelle und Marktmodelle". In: *Zeitschrift für Pädagogik* 41, S. 55–72.
- (2006). *Neue Theorie der Schule: Einführung in das Verstehen von Bildungssystemen*. Wiesbaden: VS Verlag für Sozialwissenschaften.
- (2008). *Schule gestalten: Systemsteuerung, Schulentwicklung und Unterrichtsqualität*. Wiesbaden: VS Verlag für Sozialwissenschaften. (Besucht am 26. 09. 2014).
- (2014). "Die Wirksamkeit der neuen Steuerung: Theoretische und methodische Probleme ihrer Steuerung". In: *Educational Governance als Forschungsperspektive*. Hrsg. von Katharina Maag Merki, Roman Langer und Herbert Altrichter. Educational Governance. Wiesbaden: VS Verlag für Sozialwissenschaften, S. 27–50.
- Feuser, Georg (1995). *Behinderte Kinder und Jugendliche: Zwischen Integration und Aussonderung*. Darmstadt: Wissenschaftliche Buchgesellschaft.
- (1999). "Integration - eine Frage der Didaktik einer Allgemeinen Pädagogik". In: *Behinderte in Familie, Schule und Gesellschaft* 1, S. 39–49.
- Franzen, Martina (2015). "Der Impact Faktor war gestern". In: *Soziale Welt* 66.2, S. 225–242.
- Füller, Christian (2009). *Die gute Schule: Wo unsere Kinder gerne lernen*. München: Patloch Verlag.
- Füssel, Hans-Peter (2008). "Schulinspektion und Schulaufsicht, Schulinspektion oder Schulaufsicht, Schulinspektion versus Schulaufsicht, Schulinspektion als Schulaufsicht". In: *Externe Evaluation von Schulen*. Hrsg. von Hans Döbert und Kathrin Dederling. Münster, New York, München und Berlin: Waxmann, S. 153–164.
- Gärtner, Holger (2013). "Praxis und Verhältnis interner und externer Evaluation im Schulsystem im internationalen Vergleich". In: *Zeitschrift für Erziehungswissenschaft* 16.4, S. 693–712.
- Gärtner, Holger, Hüsemann, Denis und Pant, Hans Anand (2009). "Wirkungen von Schulinspektion aus Sicht betroffener Schulleitungen. Die Brandenburger Schulleiterbefragung". In: *Empirische Pädagogik* 23.1, S. 1–18.
- Gärtner, Holger und Pant, Hans Anand (2011). "How valid are school inspections? Problems and strategies for validating processes and results". In: *Studies in Educational Evaluation* 37.2-3.
- Gärtner, Holger, Wurster, Sebastian und Pant, Hans Anand (2013). "The effect of school inspections on school improvement". In: *School Effectiveness and School Improvement* 25.4, S. 489–508.
- Gasterstädt, Julia (2016). "Querschnittsthema oder Aufgabe neben anderen? Inklusion als Innovationsaufforderung in Bildungssystemen". In: *Institutioneller Wandel im Bildungswesen*. Hrsg. von Björn Hermstein, Nils Berkemeyer und Veronika Manitius. Institutionenforschung im Bildungsbereich. Weinheim und Basel: Beltz Juventa, S. 72–90.
- Gasterstädt, Julia und Urban, Michael (2016). "Einstellung zu Inklusion? Implikationen aus Sicht qualitativer Forschung im Kontext der Entwicklung inklusiver Schulen". In: *Empirische Sonderpädagogik* 1, S. 54–66.
- Gebel, Tobias, Grenzer, Matthis, Kreuzsch, Julia, Liebig, Stefan, Schuster, Heidi, Tschewinka, Ralf, Watteler, Oliver und Witzel, Andreas (2015). "Verboten ist, was nicht ausdrücklich erlaubt ist: Datenschutz in qualitativen Interviews". In: *Forum Qualitative Sozialforschung* 16.2. URL: <http://www.qualitative-research.net/index.php/fqs/rt/printerFriendly/2266/3821> (besucht am 03. 04. 2017).
- GEW Bayern (2018). *Externe Evaluation an Schulen wird ausgesetzt*. URL: <https://bildungsklick.de/schule/meldung/externe-evaluation-an-schulen-wird-ausgesetzt/> (besucht am 22. 11. 2018).
- Gläser, Jochen und Laudel, Grit (2009). *Experteninterviews und qualitative Inhaltsanalyse als Instrumente rekonstruierender Untersuchungen*. Lehrbuch. Wiesbaden: VS Verlag für Sozialwissenschaften.
- Goffman, Erving (2002). *Stigma: Über Techniken der Bewältigung beschädigter Identität*. Bd. 140. Suhrkamp-Taschenbuch Wissenschaft. Frankfurt am Main: Suhrkamp.
- Gomolla, Mechthild (2009). "Leistungsbeurteilung in der Schule: Zwischen Selektion und Förderung, Gerechtigkeitsanspruch und Diskriminierung". In: *Migration und schulischer Wandel: Elternbeteiligung*. Hrsg. von Sara Fürstenau und Mechthild Gomolla. Wiesbaden: VS Verlag für Sozialwissenschaften, S. 25–50.
- Gomolla, Mechthild und Radtke, Frank-Olaf (2002). *Institutionelle Diskriminierung*. Wiesbaden: VS Verlag für Sozialwissenschaften.
- Göransson, Kerstin und Nilholm, Claes (2014). "Conceptual diversities and empirical shortcomings – a critical analysis of research on inclusive education". In: *European Journal of Special Needs Education* 29.3, S. 265–280. URL: <http://www.tandfonline.com/doi/pdf/10.1080/08856257.2014.933545> (besucht am 17. 07. 2015).

- Görtler, Sophie (2015). "Indikatoren für inklusive Schulentwicklung. Eine qualitative Inhaltsanalyse vorhandener Qualitätsindikatoren: Unveröffentlichte Masterarbeit an der Humboldt Universität". Diss. Berlin: Humboldt Universität zu Berlin.
- Gorur, Radhika (2014). "Towards a Sociology of Measurement in Education Policy". In: *European Educational Research Journal* 13.1, S. 58–72.
- Graumann, Sigrid (2012). "Inklusion geht weit über „Dabeisein“ hinaus – Überlegungen zur Umsetzung der UN-Behindertenrechtskonvention in der Pädagogik". In: *Soziale Inklusion*. Hrsg. von Hans-Jürgen Balz, Benjamin Benz und Carola Kuhlmann. Wiesbaden: Springer VS, S. 79–93.
- Gresch, Cornelia und Piezunka, Anne (2015). "Schülerinnen und Schüler mit sonderpädagogischer Förderung (im Bereich „Lernen“) an Regelschulen". In: *Inklusion von Schülerinnen und Schülern mit sonderpädagogischem Förderbedarf in Schulleistungserhebungen*. Hrsg. von Poldi Kuhl, Petra Stanat, Birgit Lütje-Klose, Cornelia Gresch, Hans Anand Pant und Manfred Prenzel. Wiesbaden: Springer VS, S. 181–219.
- Grosche, Michael (2015). "Was ist Inklusion? Ein Diskussions- und Positionsartikel zur Definition von Inklusion aus Sicht der empirischen Bildungsforschung". In: *Inklusion von Schülerinnen und Schülern mit sonderpädagogischem Förderbedarf in Schulleistungserhebungen*. Hrsg. von Poldi Kuhl, Petra Stanat, Birgit Lütje-Klose, Cornelia Gresch, Hans Anand Pant und Manfred Prenzel. Wiesbaden: Springer VS, S. 17–39.
- Grundmann, Gunhild, Körtters, Catrin und Krüger, Heinz-Hermann (2000). "Schulische Partizipationsstrukturen und -möglichkeiten". In: *Jugendliche Lebenswelten und Schulentwicklung*. Hrsg. von Heinz-Hermann Krüger, Gunhild Grundmann, Catrin Körtters und Nicolle Pfaff. Studien zur Schul- und Bildungsforschung. Wiesbaden und s.l.: VS Verlag für Sozialwissenschaften, S. 257–274.
- Grundmann, Gunhild und Kramer, Rolf-Torsten (2001). "Partizipation als schulische Dimension - Demokratische Reformhoffnungen zwischen schulischen Gestaltungsmöglichkeiten und strukturellen Brechungen". In: *Partizipation in der Schule*. Hrsg. von Jeanette Böhme und Rolf-Torsten Kramer. Opladen: Leske + Budrich, S. 59–92.
- Hadjar, Andreas, Hrsg. (2011). *Geschlechtsspezifische Bildungsungleichheiten*. Wiesbaden: VS Verlag für Sozialwissenschaften.
- Halverson, Richard, Kelley, Carolin und Kimball, Steven (2004). "Implementing Teacher Evaluation Systems: How principals make sense of complex artifacts to shape local instructional practice". In: *Educational Administration, Policy and Reform* 3, S. 153–188.
- Hamann, Julian (2016). "Peer Review post mortem. Bewertungen in akademischen Nachrufen". In: *Berliner Journal für Soziologie* 26.3–4, S. 433–457.
- Hartmann, Jutta (2012). *Institutionen, die unsere Existenz bestimmen: Heteronormativität und Schule*. Hrsg. von Bundeszentrale für politische Bildung. URL: <http://www.bpb.de/apuz/150624/heteronormativitaet-und-schule?p=all> (besucht am 12. 04. 2014).
- Hartung-Beck, Viola und Diemer, Tobias (2009). "Sensemaking durch Outputorientierung: Erfahrungen mit der Nutzung von Lernstandserhebungen in Schule". In: *Organisation und Erfahrung*. Hrsg. von Michael Göhlich, Susanne Maria Weber und Stephan Wolff. Wiesbaden: VS Verlag für Sozialwissenschaften, S. 215–226.
- Heid, Helmut (2000). "Qualität. Überlegungen zur Begründung einer pädagogischen Beurteilungskategorie". In: *Zeitschrift für Pädagogik* 41, S. 41–51.
- (2007). "Qualität von Schule - Zwischen Bildungstheorie und Bildungsökonomie". In: *Qualität von Schule*. Hrsg. von Jürgen van Buer und Cornelia Wagner. Frankfurt am Main: Peter Lang, S. 55–66.
- (2013). "Logik, Struktur und Prozess der Qualitätsbeurteilung von Schule und Unterricht". In: *Zeitschrift für Erziehungswissenschaft* 16.2, S. 405–431.
- Heinrich, Martin (2007). *Governance in der Schulentwicklung: Von der Autonomie zur evaluationsbasierten Steuerung*. Bd. 3. Educational Governance. Wiesbaden: VS Verlag für Sozialwissenschaften.
- (2010). "Bildungsgerechtigkeit durch Evidence-based-Policy? Governanceanalysen zu einem bildungspolitischen Programm". In: *Evaluation, Bildung und Gesellschaft*. Hrsg. von Wolfgang Böttcher, Jan Nikolas Dicke und Nina Hogrebe. Münster, New York, NY, München, Berlin: Waxmann, S. 47–68.
- (2015). "Zur Ambivalenz der Idee evidenzbasierter Schulentwicklung: Das Beispiel Schulinspektion - fortschrittlicher Rückschritt oder Innovation". In: *Zeitschrift für Pädagogik* 61.6, S. 778–792.
- Heinrich, Martin und Lambrecht, Maike (2016). "Fusion von School-Effectiveness- und School-Development-Research? Eine Auseinandersetzung am Beispiel der Schulinspektion". In: *Schulqualität - Bilanz und Perspektiven*. Hrsg. von Ulrich Steffens und Tino Bargel. Beiträge zur Schulentwicklung. Münster u. a.: Waxmann, S. 183–200.
- Heinrich, Martin, Lambrecht, Maike, Böhm-Kasper, Oliver, Brüsemeister, Thomas und Wissinger, Jochen (2014). "Funktionen von Schulinspektionen? Zum Governance-Programm der Vergewisserung und Weiterentwicklung der Qualität schulischer Arbeit". In: *Damit Unterricht gelingt*. Hrsg. von Christian Fischer. Münstersche Gespräche zur Pädagogik. Münster u. a.: Waxmann, S. 19–51.

- Heinrich, Martin, Urban, Michael und Werning, Rolf (2013). "Grundlagen, Handlungsstrategien und Forschungsperspektiven für die Ausbildung und Professionalisierung von Fachkräften für inklusive Schule". In: *Inklusive Bildung professionell gestalten*. Hrsg. von Hans Döbert und Horst Weishaupt. Münster u. a.: Waxmann, S. 69–133.
- Heintz, Bettina (2010). "Numerische Differenz. Überlegungen zu einer Soziologie des (quantitativen) Vergleichs". In: *Zeitschrift für Soziologie* 39.3, S. 162–181.
- Helbig, Marcel und Jähnen, Stefanie (2013). "Bildungsbenachteiligung durch Übergewicht: Warum adipöse Kinder in der Schule schlechter abschneiden". In: *Zeitschrift für Soziologie* 42.13, S. 403–423.
- Helmke, Andreas, Hornstein, Walter und Terhart, Ewald (2018). "Qualität und Qualitätssicherung im Bildungsbereich: Zur Einleitung in das Beiheft". In: *Zeitschrift für Pädagogik* (), S. 7–14. (Besucht am 10. 07. 2018).
- Helsper, Werner, Jeanette Böhme, Rolf-Torsten Kramer und Angelika Lingkost, Hrsg. (2001). *Schulkultur und Schulmythos*. Studien zur Schul- und Bildungsforschung. VS Verlag für Sozialwissenschaften.
- Helsper, Werner, Sandring, Sabine und Wiezorek, Christine (2005). "Anerkennung in pädagogischen Beziehungen Ein Problemaufriss". In: *Integrationspotenziale einer modernen Gesellschaft*. Hrsg. von Wilhelm Heitmeyer und Peter Imbusch. Analysen zu gesellschaftlicher Integration und Desintegration. VS Verlag für Sozialwissenschaften, S. 179–206.
- Hessisches Kultusministerium (2011). *Hessischer Referenzrahmen Schulqualität*.
- Hinz, Andreas (2002). "Von der Integration zur Inklusion - terminologisches Spiel oder konzeptionelle Weiterentwicklung?" In: *Zeitschrift für Heilpädagogik* 53, S. 354–361.
- (2004). "Vom sonderpädagogischen Verständnis der Integration zum integrationspädagogischen Verständnis der Inklusion!" In: *Inklusive Pädagogik*. Hrsg. von Irmtraud Schnell und Alfred Sander. Bad Heilbrunn: Verlag Julius Klinkhardt, S. 41–74.
- (2007). "Inklusion - Vision und Realität". In: *Vielfalt braucht Struktur*. Hrsg. von Dieter Katzenbach. Frankfurt am Main: Johann-Wolfgang-Goethe-Universität, S. 81–98.
- (2010). *Aktuelle Erträge der Debatte um Inklusion - worin besteht der Mehrwert gegenüber Integration?*
- Hinz, Andreas und Boban, Ines (2017). "Inklusion zwischen Menschenrechten und Neoliberalismus - eine Problem-skizze". In: *Leistung inklusive? Inklusion in der Leistungsgesellschaft*. Hrsg. von Birgit Lütje-Klose, Mai-Anh Boger, Benedikt Hopmann und Philip Neumann. Bad Heilbrunn: Verlag Julius Klinkhardt, S. 39–47.
- Hirschauer, Stefan (2015). "How Editors Decide. Oral Communication in Journal Peer Review". In: *Human Studies* 38.1, S. 37–55.
- Holtappels, Heinz Günter (2003). *Schulqualität durch Schulentwicklung und Evaluation: Konzepte, Forschungsbefunde, Instrumente*. 1. Aufl. Beiträge zur Schulentwicklung. München: Luchterhand.
- Honneth, Axel (1992). *Kampf um Anerkennung: Zur moralischen Grammatik sozialer Konflikte*. 1. Aufl. Frankfurt am Main: Suhrkamp.
- Huber, Franz (2008). "Konsequenzen aus der externen Evaluation an Bayerns Schulen". In: *Bildungsmonitoring und Bildungscontrolling in nationaler und internationaler Perspektive*. Hrsg. von Wolfgang Böttcher, Wilfried Bos, Hans Döbert und Heinz Günter Holtappels. Münster u. a.: Waxmann, S. 265–278.
- Hummel, Myriam (2016). "Die Entwicklung inklusiver Bildung in Malawi: zwischen makro-politischer Deklaration und lokaler Umsetzung". In: *Zeitschrift für Inklusion Online* 1. URL: <http://www.inklusion-online.net/index.php/inklusion-online/article/view/343/287> (besucht am 01. 12. 2018).
- Hupka-Brunner, Sandra, Imdorf, Christian, Gunder, Hans-Ulrich und Bergman, Manfred Max (2015). "Zur Einführung. Der Qualitätsdiskurs im Bildungsbereich eine kritische Bestandsaufnahme". In: *Qualität in der Bildung*. Hrsg. von Sandra Hupka-Brunner, Hans-Ulrich Grunder, Manfred Max Bergman und Christian Imdorf. Bad Heilbrunn: Verlag Julius Klinkhardt, S. 7–12.
- Husfeldt, Vera (2011). "Wirkungen und Wirksamkeit der externen Schulevaluation". In: *Zeitschrift für Erziehungswissenschaft* 14.2, S. 259–282.
- Ispa-Landa, Simone (2013). "Gender, Race, and Justifications for Group Exclusion: Urban Black Students Bussed to Affluent Suburban Schools". In: *Sociology of Education* 86.3, S. 218–233. (Besucht am 06. 10. 2016).
- Jarke, Juliane und Breiter, Andreas (2016). *Datafying education: How digital data practices reconfigure the organisation of learning in schools*. Vortrag am WZB.
- Jeacle, Ingrid und Carter, Chris (2011). "In TripAdvisor we trust: Rankings, calculative regimes and abstract systems". In: *Accounting, Organizations and Society* 36.4-5, S. 293–309.
- Katzenbach, Dieter (2015). "Zu den Theoriefundamenten der Inklusion. Eine Einladung zum Diskurs aus der Perspektive der kritischen Theorie". In: *Herausforderung Inklusion*. Hrsg. von Irmtraud Schnell. Bad Heilbrunn: Verlag Julius Klinkhardt, S. 19–32.
- Klemm, Klaus (2014). "Bildungskatastrophe: Hat Georg Picht recht behalten?" In: *Zeit Online*. URL: <http://www.zeit.de/2014/06/bildungskatastrophe-these-georg-picht> (besucht am 07. 08. 2017).
- Koch, Barbara (2011). *Wie gelangen Innovationen in die Schule?* Wiesbaden: VS Verlag für Sozialwissenschaften.

- Krüger, Anne K. und Reinhart, Martin (2017). "Theories of Valuation – Building Blocks for Conceptualizing Valuation between Practice and Structure". In: *Historical Social Research* 42.1, S. 263–285. (Besucht am 09. 04. 2017).
- Kruschel, Robert, Hrsg. (2017). *Menschenrechtsbasierte Bildung: Inklusive und demokratische Lern- und Erfahrungswelten im Fokus*. Lebenshilfe. Bad Heilbrunn: Verlag Julius Klinkhardt.
- Kruse, Stefan und Dederich, Kathrin (2018). "The idea of inclusion: Conceptual and empirical diversities in Germany". In: *Improving Schools* 21.1, S. 19–31.
- Kultusministerkonferenz (1997). *Grundsätzliche Überlegungen zu Leistungsvergleichen innerhalb der Bundesrepublik Deutschland: Konstanzer Beschluss*. URL: http://www.kmk.org/fileadmin/veroeffentlichungen_beschluesse/1997/1997_10_24-Konstanzer-Beschluss.pdf (besucht am 19. 06. 2015).
- Kurth-Buchholz, Elke (2011). *Schülermitbestimmung aus Sicht von Schülern und Lehrern: Eine vergleichende Untersuchung an Gymnasien in Brandenburg und Nordrhein-Westfalen*. Bd. 28. Empirische Erziehungswissenschaft. Münster u. a.: Waxmann.
- Kussau, Jürgen und Brüsemeister, Thomas (2016). "Educational Governance: Zur Analyse der Handlungskoordination im Mehrebenensystem der Schule". In: *Ambivalenzen des Ökonomischen*. Hrsg. von Martin Heinrich und Barbara Kohlstock. Educational Governance. Wiesbaden: VS Verlag für Sozialwissenschaften, S. 15–54.
- Lambrecht, Maike (2013). "Schulinspektorinnen und Schulinspektoren zwischen Schulentwicklung und bildungspolitischer Innovation". In: *Innovationen im Bildungswesen*. Hrsg. von Matthias Rürup und Inka Bormann. Educational Governance. Wiesbaden: VS Verlag für Sozialwissenschaften, S. 223–247.
- (2018). *Steuerung als pädagogisches Problem: Empirische Rekonstruktionen zur Interaktion in Schulinspektions-Interviews*. Bd. 16. Rekonstruktive Bildungsforschung. Wiesbaden: Springer Fachmedien Wiesbaden.
- Lambrecht, Maike und Rürup, Matthias (2012). "Bildungsforschung im Rahmen einer evidence based policy: Das Beispiel „Schulinspektion“". In: *Schul- und Unterrichtsreform durch ergebnisorientierte Steuerung*. Hrsg. von Albrecht Wacker, Uwe Maier und Jochen Wissinger. Wiesbaden: VS Verlag für Sozialwissenschaften, S. 57–77.
- Lamnek, Siegfried (1995). *Qualitative Sozialforschung: Band 2: Methoden und Techniken*. Weinheim: Beltz.
- Lamont, Michèle (2010). *How professors think: Inside the curious world of academic judgment*. 1. paperback ed. Cambridge, Mass.: Harvard University Press.
- (2012). "Toward a Comparative Sociology of Valuation and Evaluation". In: *Annual Review of Sociology* 38.1, S. 201–221.
- Lamont, Michèle, Beljean, Stefan und Clair, Matthew (2014). "What is missing? Cultural processes and causal pathways to inequality". In: *Socio-Economic Review* 12.3, S. 573–608.
- Lamont, Michèle und Huutoniemi, Katri (2011). "Comparing Customary Rules of Fairness: Evaluative Practices in Various Types of Peer Review Panels". In: *Social Knowledge in the Making*. Hrsg. von Charles Camic, Neil Gross und Michèle Lamont. Chicago: The University of Chicago Press, S. 209–232.
- Landesinstitut für Schulentwicklung Baden-Württemberg (2015). *Qualitätsrahmen zur Fremdevaluation (zweiter Durchgang) an allgemein bildenden Schulen in Baden-Württemberg*.
- Landesinstitut für Schulqualität und Lehrerbildung Sachsen-Anhalt (2013). *Qualitätsrahmen schulischer Arbeit in Sachsen-Anhalt*. URL: https://www.bildung-lsa.de/files/c2bac8db69b228cadf64ebb7acde10ec/qualitaetsrahmen_LSA_Sep_2013.pdf (besucht am 30. 04. 2016).
- Landwehr, Norbert (2011). "Thesen zur Wirkung und Wirksamkeit der externen Schulevaluation". In: *Wirkungen und Wirksamkeit der externen Schulevaluation*. Hrsg. von Carsten Quesel, Norbert Landwehr, Peter Steiner und Vera Husfeldt. Bern: hep der Bildungsverlag, S. 35–70.
- Lindmeier, christiand und Lindmeier, Bettina (2015). "Inklusion aus der Perspektive des rechtlichen und ethischen Begründungsdiskurses". In: *Zeitschrift für Erziehungswissenschaft* 26.51, S. 43–51.
- Lipsky, Michael (2010). *Street-Level Bureaucracy, Dilemmas of the Individual in Public Service*. New York: Russell Sage Foundation.
- Lohmann, A. und Reißmann, J. (2007). "Von der Verstrickung der Rollen zum funktionalen Unterstützungsnetz - Die Schulinspektion im Gefüge von Schulleitung, Kollegium, Schulaufsicht und Unterstützungssystem". In: *Journal für Schulentwicklung* 11.3, S. 15–24.
- Lowenhaupt, Rebecca, Spillane, James P. und Hallet, Tim (2016). "Education Policy in Leadership practice: Accountability talk in schools". In: *Journal of School Leadership* 26, S. 783–810.
- Luhmann, Niklas (1995). "Inklusion und Exklusion". In: *Soziologische Aufklärung* 6. Hrsg. von Niklas Luhmann. Opladen: Westdeutscher Verlag, S. 237–264.
- Lussi, Isabella und Huber, Stephan (2015). "Das Erleben von Anerkennung in der Schule und seine Relevanz für die Werteentwicklung von jungen Erwachsenen". In: *Forum Qualitative Sozialforschung* 16.3.
- Maag Merki, Katharina, Roman Langer und Herbert Altrichter, Hrsg. (2014). *Educational Governance als Forschungsperspektive: Strategien. Methoden. Ansätze*. Bd. 17. Educational Governance. Wiesbaden: Springer VS.
- Maitlis, Sally und Christianson, Marlys (2014). "Sensemaking in Organizations: Taking Stock and Moving Forward". In: *The Academy of Management Annals* 8.1, S. 57–125.

- Maritzen, Norbert (2006). "Eine Trendanalyse: Schulinspektion zwischen Aufsicht und Draufsicht". In: *Schulinspektion und Schulleitung*. Hrsg. von Herbert Buchen. Stuttgart: Raabe, S. 7–25.
- (2008). "Schulinspektionen". In: *Die deutsche Schule* 100.1, S. 85–96.
- Mayring, Philipp (2010). "Qualitative Inhaltsanalyse". In: *Handbuch Qualitative Forschung in der Psychologie*. Hrsg. von Günter Mey und Katja Mruck. Wiesbaden: VS Verlag für Sozialwissenschaften, S. 597–609.
- Mayring, Philipp und Fenzl, Thomas (2014). "Qualitative Inhaltsanalyse". In: *Handbuch Methoden der empirischen Sozialforschung*. Hrsg. von Nina Baur und Jörg Blasius. Wiesbaden: VS Verlag für Sozialwissenschaften, S. 543–556.
- Mehrpouya, Afshin und Samiolo, Rita (2016). "Performance measurement in global governance: Ranking and the politics of variability". In: *Accounting, Organizations and Society* 55, S. 12–31.
- Meier, Frank, Peetz, Thorsten und Waibel, Désirée (2016). "Bewertungskonstellationen. Theoretische Überlegungen zur Soziologie der Bewertung". In: *Berliner Journal für Soziologie* 26.3-4, S. 307–328.
- Merkus, Sander, Willems, Thijs, Schipper, Danny, van Marrewijk, Alfons, Koppenjan, Joop, Veenswijk, Marcel und Bakker, Hans (2016). "A Storm is Coming? Collective Sensemaking and Ambiguity in an Inter-organizational Team Managing Railway System Disruptions". In: *Journal of Change Management* 17.3, S. 228–248.
- Meyer, Heinz-Dieter und Zahedi, Katie (2014). *Nein zu Pisa: Offener Brief an Andreas Schleicher*. URL: <http://bildung-wissen.eu/fachbeitraege/nein-zu-pisa-offener-brief-an-andreas-schleicher.html> (besucht am 03.04.2018).
- Mills, Jean Helms, Thurlow, Amy und Mills, Albert J. (2010). "Making sense of sensemaking: The critical sensemaking approach". In: *Qualitative Research in Organizations and Management: An International Journal* 5.2, S. 182–195.
- Ministerium für Bildung, Jugend und Sport Brandenburg (2016). *Orientierungsrahmen Schulqualität im Land Brandenburg: Ein Handbuch für gute Schulen im Land Brandenburg*. (Besucht am 30.10.2017).
- Ministerium für Bildung, Jugend und Sport des Landes Brandenburg (2008). *Orientierungsrahmen Schulqualität in Brandenburg: Qualitätsbereiche und Qualitätsmerkmale guter Schulen*. URL: http://www.mbj.s.brandenburg.de/sixcms/media.php/5527/Orientierungsrahmen_Schulqualitaet_Endversion2008.pdf (besucht am 04.07.2014).
- Ministerium für Bildung, Wissenschaft, Jugend und Kultur, Rheinland-Pfalz (2015). *Agentur für Qualitätssicherung, Evaluation und Selbstständigkeit von Schulen soll aufgelöst werden*. URL: <https://bildungsklick.de/schule/meldung/agentur-fuer-qualitaetssicherung-evaluation-und-selbststaendigkeit-von-schulen-soll-aufgeloeset-werden/> (besucht am 21.07.2018).
- Ministerium für Kultus, Jugend und Sport Baden-Württemberg (2017). *Deutsche Schulen im Ausland*. URL: <http://www.km-bw.de/Lde/Startseite/Schule/Deutsche+Schulen+im+Ausland> (besucht am 29.10.2017).
- Ministerium für Schule und Berufsbildung Schleswig-Holstein (2016). *Handbuch: Schulfeedback.SF*. (Besucht am 24.10.2017).
- Moldenhauer, Anna (2015). *Dialektik der Partizipation: Eine rekonstruktive Studie zu Erfahrungen von Schülerinnen und Schülern mit Partizipation in Gemeinschaftsschulen*. Studien zur Bildungsgangforschung. Opladen, Berlin und Toronto: Verlag Barbara Budrich.
- Moser, Vera (2012a). "Braucht die Inklusionspädagogik einen Behinderungsbegriff". In: *Zeitschrift für Inklusion Online* 3. URL: <http://www.inklusion-online.net/index.php/inklusion-online/article/view/40/40> (besucht am 29.01.2014).
- (2012b). "Inklusion: Standards und Perspektiven". In: *Sonderpädagogik in Berlin* 2, S. 23–27.
- (2017). "Inklusion und Organisationsentwicklung: Eine Einführung". In: *Schulentwicklung und Inklusion*. Hrsg. von Vera Moser und Marina Egger. Stuttgart: W. Kohlhammer, S. 9–31.
- Moser, Vera und Marina Egger, Hrsg. (2017). *Schulentwicklung und Inklusion*. Stuttgart: W. Kohlhammer.
- Neißl, Marianne (2016). "Inklusion - eine streitbare Leitidee". In: *Heterogenität in pädagogischen Handlungsfeldern*. Hrsg. von Danièle Hollick, Martin Kramer, Johannes Reitingen und Marianne Neißl. Kassel: Kassel University Press, S. 1–6.
- Nesyba, Thea (2018). *Bewegung durch kollektive Akteure und ihre politischen Narrative - eine Fallstudie zur niedersächsischen Schulpolitik*. Essen.
- Netzwerk Menschenrechte (2017). *Übereinkommen über die Rechte des Kindes › UN-Kinderrechtskonvention*. URL: <https://www.kinderrechtskonvention.info/> (besucht am 04.08.2017).
- Niedersächsisches Kultusministerium (2014). *Orientierungsrahmen: Schulqualität in Niedersachsen*.
- Niedersächsisches Landesinstitut für schulische Qualitätsentwicklung (2014). *Zum Entwicklungsauftrag der Schulinspektion: Grundlagen des weiterentwickelten Inspektionsverfahrens an allgemeinbildenden Schulen*.
- (2015). *Das Kernaufgabenmodell als Instrument für das Qualitätsmanagement an allgemein bildenden Schulen*.
- Nijhof, André und Jeurissen, Ronald (2006). "Editorial: A sensemaking perspective on corporate social responsibility: introduction to the special issue". In: *Business Ethics: A European Review* 15.4, S. 316–322. (Besucht am 15.10.2017).
- O'Neil, Cathy (2018). *Do Algorithms Perpetuate Human Bias?* National Public Radio.
- Olbrisch, Miriam (2017). "Warum ein Kultusminister die Inklusion für gescheitert hält: Interview mit Marco Tullner". In: *Spiegel Online*. URL: <http://www.spiegel.de/spiegel/inklusion-sachsen-anhalt-will-fuer-foerderschulen-kaempfen-a-1184740.html> (besucht am 21.06.2018).

- Osley, Audrey und Morrison, Marlene (2000). *Inspecting Schools for Race Equality: OFSTED' strengths and weaknesses*. Hrsg. von Executive Summary of a Report for the Commission of Racial Equality.
- Penninckx, Maarten (2016). "Effects and side effects of school inspections: A general framework". In: *Studies in Educational Evaluation* 52, S. 1–11.
- Penuel, William R., Farrell, Caitlin C., Allen, Anna-Ruth, Toyama, Yukie und Coburn, Cynthia E. (2018). "What Research District Leaders Find Useful". In: *Educational Policy* 32.4, S. 540–569.
- Pfahl, Lisa (2011). *Techniken der Behinderung: Der deutsche Lernbehinderungsdiskurs, die Sonderschule und ihre Auswirkungen auf Bildungsbiografien*. Freie Univ., Diss.–Berlin, 2009. Bd. 7. Disability Studies. Bielefeld: Transcript Verlag.
- Picht, Georg (1964). *Die deutsche Bildungskatastrophe*. Olten: Walter-Verlag.
- Piezunka, Anne (2015). *Das bunte Klassenzimmer – welches Handwerkszeug brauchen Lehrkräfte im Umgang mit Heterogenität?* Hrsg. von Bundeszentrale für politische Bildung. URL: <http://www.bpb.de/gesellschaft/kultur/zukunftsbildung/185859/umgang-mit-heterogenitaet?p=all> (besucht am 16. 08. 2016).
- (2017). "Ist eine „gute“ Schule eine „inklusive“ Schule? Bewertungskriterien von Schulinspektionen in Deutschland (Hamburg/ Sachsen-Anhalt) und England". In: *Schulentwicklung und Inklusion*. Hrsg. von Vera Moser und Marina Egger. Stuttgart: W. Kohlhammer, S. 162–180.
- (2019). "Struggle for acceptance: Maintaining external school inspection as an institution". In: *Historical Social Research* 44.2.
- Piezunka, Anne, Gresch, Cornelia, Sälzer, Christine und Kroth, Anna (2016). "Identifizierung von Schülerinnen und Schülern nach Vorgaben der UN-BRK in bundesweiten Erhebungen: Sonderpädagogischer Förderbedarf, sonderpädagogische Förderung oder besondere Unterstützung?" In: *Zeitschrift für Pädagogik* 62, S. 190–211.
- Piezunka, Anne, Gresch, Cornelia und Wrase, Michael (2018). "Artikel 24 der UN-Behindertenrechtskonvention: Möglichkeiten und Grenzen beim Monitoring von Menschenrechten mit Blick auf die Einzelschulebene". In: *Zeitschrift für Inklusion Online* 1. URL: <https://www.inklusion-online.net/index.php/inklusion-online/article/view/369> (besucht am 03. 12. 2018).
- Piezunka, Anne, Schaffus, Tina und Grosche, Michael (2017). "Vier Verständnisse von schulischer Inklusion und ihr gemeinsamer Kern: Ergebnisse von Experteninterviews mit Inklusionsforscherinnen und -forschern". In: *Unterrichtswissenschaft* 4, S. 207–222.
- Platte, Andrea (2014). "Inklusive Bildungsprozesse - Teilhaben am Lernen und Lehren in einer Schule für alle". In: *Teilhaben an Schule*. Hrsg. von Thomas Rihm. Wiesbaden: Springer VS, S. 73–86.
- Powell, Justin J. W. (2006). *Barriers to Inclusion: Special Education in the United States and Germany*. Florence: Taylor and Francis.
- (2007). "Behinderung in der Schule, behindert durch Schule? Die Institutionalisierung der <schulischen Behinderung>". In: *Disability Studies, Kultursoziologie und Soziologie der Behinderung*. Hrsg. von Anne Waldschmidt und Werner Schneider. Bielefeld: Transcript Verlag, S. 321–343.
- Power, Michael (1997). *The audit society: Rituals of verification*. Oxford und New York: Oxford University Press.
- Prenzel, Annedore (2001). "Egalitäre Differenz in der Bildung". In: *Unterschiedlich verschieden*. Hrsg. von Helma Lutz und Norbert Wenning. Opladen: Leske + Budrich, S. 93–108.
- (2003). "Kinder akzeptieren, diagnostizieren, etikettieren? Kulturen- und Leistungsvielfalt im Bildungswesen". In: *Heterogenität macht Schule*. Hrsg. von Birgit Warzecha. Novemberakademie. Münster u. a.: Waxmann, S. 27–40.
- (2006). *Pädagogik der Vielfalt: Verschiedenheit und Gleichberechtigung in interkultureller, feministischer und integrativer Pädagogik*. 3. Aufl. Wiesbaden: VS Verlag für Sozialwissenschaften.
- (2013). *Pädagogische Beziehungen zwischen Anerkennung, Verletzung und Ambivalenz*. Leverkusen: Verlag Barbara Budrich.
- (2014). "Heterogenität oder Lesarten von Gleichheit und Freiheit in der Bildung." In: *Heterogenität - Zur Konjunktur eines pädagogischen Konzepts*. Hrsg. von Hans-Christoph Koller, Rita Casale und Norbert Ricken. Schriftenreihe der Kommission Bildungs- und Erziehungsphilosophie in der Deutschen Gesellschaft für Erziehungswissenschaft. Paderborn: Ferdinand Schöningh, S. 45–67.
- Prenzel, Annedore und Ursula Winklhofer, Hrsg. (2014). *Kinderrechte in pädagogischen Beziehungen: Band 2: Forschungszugänge*. Opladen, Berlin und Toronto: Verlag Barbara Budrich.
- Prenzel, Manfred, Baumert, Jürgen, Blum, Werner, Lehmann, Rainer, Leutner, Detlev, Neubrand, Michael, Pekrun, Reinhard, Rost, Jürgen und Schiefele, Ulrich (2005). *PISA 2003: Ergebnisse des zweiten Ländervergleichs: Zusammenfassung*. Hrsg. von PISA-Konsortium Deutschland.
- Projektbüro Inklusion (2014). *Checkliste Inklusion*. Hrsg. von Institut für Qualitätsentwicklung in Hessen. URL: www.inklusion-schule.info/pdf/checkliste-inklusion.pdf (besucht am 08. 10. 2014).
- Quinn, John (2009). "Teacher Sense-making and policy implementation: A qualitative case study of a school district's reading initiative in science". Dissertation. (Besucht am 01. 01. 2016).

- Radtke, Frank-Olaf und Gomolla, Mechthild (2009). *Institutionelle Diskriminierung: Die Herstellung ethnischer Differenz in der Schule*. Wiesbaden: VS Verlag für Sozialwissenschaften.
- Reitz, Sandra (2015). *Kinder und Jugendliche haben ein Recht auf Partizipation: Was aus menschenrechtlicher Sicht im Bildungsbereich getan werden muss*. Hrsg. von Deutsches Institut für Menschenrechte. URL: http://www.institut-fuer-menschenrechte.de/fileadmin/user_upload/Publikationen/Policy_Paper/PP_31__Kinder_und_Jugendliche_haben_ein_Recht_auf_Partizipation.pdf (besucht am 30. 09. 2015).
- Ribolits, Erich (2009). "Bildungsqualität - was ist das und woher rührt die grassierende Sorge um dieselbe?" In: *Bildungsqualität!* Hrsg. von Eveline Christof, Erich Ribolits und Johannes Zuber. Schulheft. Innsbruck: Studienverlag, S. 7–18.
- Robert Bosch Stiftung (2018). *Der Deutsche Schulpreis*. URL: <http://schulpreis.bosch-stiftung.de/content/language1/html/index.asp> (besucht am 10. 07. 2018).
- Rottenburg, Richard und Merry, Sally E. (2015a). "A world of indicators: the making of governmental knowledge through quantification". In: *The world of indicators*. Hrsg. von Richard Rottenburg, Sally E. Merry, Sung-Joon Park und Johanna Mugler. Cambridge: Cambridge University Press, S. 1–33.
- (2015b). "A world of indicators: The making of governmental knowledge through Quantification". In: *The world of indicators*. Hrsg. von Richard Rottenburg. Cambridge studies in law and society. Cambridge: Cambridge University Press, S. 1–33.
- Rüb, Friedrich W. und Straßheim, Holger (2012). "Politische Evidenz - Rechtfertigung durch Verobjektivierung". In: *Der Aufstieg der Legitimitätspolitik*. Hrsg. von Anna Geis, Frank Nullmeier und Christopher Daase. Leviathan : Sonderband. Baden-Baden: Nomos, S. 377–397.
- Rudolf, Beate (2017). "Teilhabe als Menschenrecht - eine grundlegende Betrachtung". In: *Teilhabe für alle ?!* Hrsg. von Elke Diehl. Bonn: Bundeszentrale für politische Bildung, S. 13–43.
- Rürup, Matthias (2008). "Typen der Schulinspektion in den deutschen Bundesländern". In: *Die deutsche Schule*, S. 467–477.
- Salais, Robert (2006). "Reforming the European Social Model and the politics of indicators: from the unemployment rate to the employment rate in the European Employment Strategy". In: *Unwrapping the European Social Model*. Hrsg. von Maria Jepsen und Amparo Serrano Pascual. Policy Press: University of Bristol, S. 189–212.
- (2008). "Evaluation und Politik: Auf der Suche nach guten Indikatoren für die Forschung". In: *Wissenschaft unter Beobachtung: Sonderheft Leviathan*. Hrsg. von Hildegard Matthies und Dagmar Simon. Wiesbaden: VS Verlag für Sozialwissenschaften, S. 193–212.
- Sandberg, Jörgen und Tsoukas, Haridimos (2015). "Making sense of the sensemaking perspective: Its constituents, limitations, and opportunities for further development". In: *Journal of Organizational Behavior* 36.S1, S6–S32.
- Sandring, Sabine (2013). *Schulversagen und Anerkennung: Scheiterende Schulkarrieren im Spiegel der Anerkennungsbedürfnisse Jugendlicher*. Wiesbaden: VS Verlag für Sozialwissenschaften.
- Sauder, Michael und Espeland, Wendy Nelson (2009). "The Discipline of Rankings: Tight Coupling and Organizational Change". In: *American Sociological Review* 74.1, S. 63–82.
- Schäfer, Hilmar (2016). "„Outstanding universal value“. Die Arbeit an der Universalisierung des Wertvollen im UNESCO-Welterbe". In: *Berliner Journal für Soziologie* 26.3–4, S. 353–375.
- Scharenberg, Katja (2015). "Schulische Kontexteffekte im erziehungswissenschaftlichen Qualitäts- und Forschungsdiskurs". In: *Qualität in der Bildung*. Hrsg. von Sandra Hupka-Brunner, Hans-Ulrich Grunder, Manfred Max Bergman und Christian Imdorf. Bad Heilbrunn: Verlag Julius Klinkhardt, S. 51–74.
- Schreier, Magrit (2014). "Varianten qualitativer Inhaltsanalyse: Ein Wegweiser im Dickicht der Begrifflichkeiten". In: *Forum Qualitative Sozialforschung* 15.1. (Besucht am 12. 02. 2015).
- Schwartz, Heather, Hamilton, Laura, Stecher, Brian und Steele, Jennifer (2016). *Expanded Measures of School Performance*. URL: http://www.rand.org/content/dam/rand/pubs/technical_reports/2011/RAND_TR968.pdf (besucht am 06. 10. 2016).
- Scott, Richard (2008). *Institutions and Organizations: Ideas and Interests*. Thousand Oaks: Sage Publications.
- Sellar, Sam (2014). "A feel for numbers: Affect, data and education policy". In: *Critical Studies in Education* 56.1, S. 131–146.
- Senatsverwaltung für Bildung, Jugend und Wissenschaft (2012). *Handlungsrahmen Schulqualität in Berlin*. (Besucht am 07. 01. 2014).
- Simmel, Georg (1900). *Philosophie des Geldes*. Köln: Anaconda Verlag.
- Simon, Jacqueline und Simon, Toni (2013). "Inklusive Diagnostik – Wesenszüge und Abgrenzung von traditionellen "Grundkonzepten" diagnostischer Praxis. Eine Diskussionsgrundlage". In: *Zeitschrift für Inklusion Online* 4. URL: <http://www.inklusion-online.net/index.php/inklusion-online/article/view/194/200> (besucht am 15. 04. 2014).
- Solga, Heike und Rolf Becker, Hrsg. (2012). *Soziologische Bildungsforschung*. Bd. 52. KZfSS - Kölner Zeitschrift für Soziologie und Sozialpsychologie. Wiesbaden: VS Verlag für Sozialwissenschaften.

- Solga, Heike und Dombrowski, Rosine (2009). *Soziale Ungleichheiten in schulischer und außerschulischer Bildung: Stand der Forschung und Forschungsbedarf*. Hrsg. von Arbeitspapier 171. Berlin: Hans-Böckler-Stiftung. URL: http://www.boeckler.de/pdf/p_arbp_171.pdf (besucht am 28.05.2012).
- Sowada, Moritz G. und Dederich, Kathrin (2014). "Ermessensspielräume in der Bewertungsarbeit von Schulinspektor/innen". In: *Zeitschrift für Bildungsforschung* 4.2, S. 119–135.
- (2016). "Die Reform der Reform". In: *Schulinspektion als Steuerungsimpuls?* Hrsg. von Arbeitsgruppe Schulinspektion. Wiesbaden: VS Verlag für Sozialwissenschaften, S. 169–199.
- Spillane, James P. (2004). *Standards Deviation: How schools Misunderstood Education Policy*. Cambridge: Harvard University Press.
- Spillane, James P., Reiser, Brian und Reimer, Todd (2002). "Policy Implementation and Cognition: Reframing and Refocusing Implementation Research". In: *Review of Educational Research* 72.3, S. 387–431.
- Sprietsma, Maresa (2009). *Discrimination in Grading? Experimental Evidence from Primary School*. URL: <ftp://ftp.zew.de/pub/zew-docs/dp/dp09074.pdf> (besucht am 26.03.2012).
- Stanat, Petra, Artelt, Cordula, Baumert, Jürgen, Klieme, Eckhard, Neubrand, Michael, Prenzel, Manfred, Schiefele, Ulrich, Schneider, Wolfgang, Tillmann, Klaus-Jürgen und Weiß, Manfred (2002). *PISA 2000: Die Studie im Überblick. Grundlagen, Methoden und Ergebnisse*. Max Planck Institut für Bildungsforschung.
- Stark, David (2009). *The sense of dissonance: Accounts of worth in economic life*. Princeton: Princeton University Press.
- Statistisches Bundesamt (2018). *Migrationshintergrund*. URL: <https://www.destatis.de/DE/ZahlenFakten/GesellschaftStaat/Bevoelkerung/MigrationIntegration/Glossar/Migrationshintergrund.html> (besucht am 28.07.2018).
- SteBis (2017). *Forschungsschwerpunkt Steuerung im Bildungssystem*. URL: <http://www.stebis.de/index.html> (besucht am 27.06.2018).
- Straßheim, Holger (2015). "Politics and policy expertise: towards a political epistemology". In: *Handbook of critical policy studies*. Hrsg. von Anna Durnová, Douglas Torgerson, Frank Fischer und Michael Orsini. Handbooks of research on public policy. Cheltenham und Northampton: Edward Elgar Publishing, S. 319–340.
- Strauss, Valerie (2017). "What to know before using school ratings tools from real estate companies". In: *The Washington Post*. URL: https://www.washingtonpost.com/news/answer-sheet/wp/2017/06/14/what-to-know-before-using-school-ratings-tools-from-real-estate-companies/?noredirect=on&utm_term=.56f939f4448a (besucht am 10.07.2018).
- Streib-Brzić, Uli und Quadflieg, Christiane (2011). *School is out?! Vergleichende Studie <Erfahrungen von Kindern aus Regenbogenfamilien in der Schule> durchgeführt in Deutschland, Schweden und Slowenien*.
- Stryck, Tom (2000). "Qualitätssicherung in der Geisterbahn: Was hat Schulaufsicht mit Schulqualität zu tun?" In: *Qualität und Qualitätssicherung im Bildungsbereich*. Hrsg. von Andreas Helmke, Walter Hornstein und Ewald Terhart. Weinheim: Beltz, S. 111–125.
- Supik, Linda (2017). "Statistik und Diskriminierung". In: *Handbuch Diskriminierung*. Hrsg. von Aladin El-Mafaalani und Gökçen Yüksel. Springer Reference Sozialwissenschaften. Wiesbaden: Springer Fachmedien Wiesbaden, S. 191–207.
- Tellisch, Christin (2016). "Inklusion braucht Kommunikation - Impulse für gelingende Inklusion in der Schule durch reflektierte, anerkennende Lehrer-Schüler-Interaktionen". In: *Zeitschrift für Inklusion Online* 1. URL: <https://www.inklusion-online.net/index.php/inklusion-online/article/download/339/283?inline=1> (besucht am 21.11.2017).
- Terhart, Ewald (2000). "Qualität und Qualitätssicherung im Schulsystem. Hintergründe - Konzepte - Probleme". In: *Zeitschrift für Pädagogik* 46, S. 809–829.
- Thévenot, Laurent (2009). "Governing Life by Standards: A View from Engagements". In: *Social Studies of Science* 39.5, S. 793–813.
- Timmermans, Stefan und Epstein, Steven (2010). "A World of Standards but not a Standard World: Toward a Sociology of Standards and Standardization". In: *Annual Review of Sociology* 36.1, S. 69–89.
- Trautmann, Matthias und Wischer, Beate (2009). "Das Konzept der Inneren Differenzierung - eine vergleichende Analyse der Diskussion der 1970er Jahre mit dem aktuellen Heterogenitätsdiskurs". In: *Perspektiven der Didaktik*. Hrsg. von Meinert Meyer, Manfred Prenzel und Stephanie Hellekamps. Wiesbaden: VS Verlag für Sozialwissenschaften, S. 159–172.
- UNESCO (2017a). *Sustainable Development Goal 4: Ensure inclusive and equitable quality education and promote lifelong learning opportunities for all*. URL: <https://sustainabledevelopment.un.org/sdg4> (besucht am 30.07.2017).
- (2017b). *What does inclusive education mean?* URL: <https://www.unesco.de/en/bildung/inklusive-bildung/hintergrund-inklusive-bildung/inklusive-bildung-inhalte.html> (besucht am 22.07.2017).
- Hrsg. (2008). *Inclusive education: the way of the future*. (Besucht am 22.07.2017).
- Verband Bildung und Erziehung e.V. (2017). *Inklusion an Schulen aus Sicht der Lehrkräfte in Deutschland - Meinungen, Einstellungen und Erfahrungen: Ergebnisse einer repräsentativen Lehrerbefragung*. Hrsg. von forsa.

- Wagner, Peter (2011). "Handlung, Institution, Kritik: Materialien zur immer noch notwendigen Erneuerung der Gesellschaftstheorie". In: *Soziologische Revue* 34.3, S. 270–276.
- Weick, Karl E. (1979). *The social psychology of organizing*. 2. ed. Topics in social psychology. New York: McGraw-Hill Inc.
- (1985). *Der Prozess des Organisierens*. Frankfurt am Main: Suhrkamp.
- (1993). "The Collapse of Sensemaking in Organizations: The Mann Gulch Disaster". In: *Administrative Science Quarterly* 38, S. 628–652.
- (1995). *Sensemaking in organizations*. Foundations for organizational science. Thousand Oaks: Sage Publications.
- Weick, Karl E., Sutcliffe, Kathleen M. und Obstfeld, David (2005). "Organizing and the Process of Sensemaking". In: *Organization Science* 16.4, S. 409–421.
- Werning, Rolf (2014). "Stichwort: Schulische Inklusion". In: *Zeitschrift für Erziehungswissenschaft*, S. 1–23.
- Wetzel, Ralf (2001). "Kognition und Sensemaking". In: *Moderne Organisationstheorien: Eine sozialwissenschaftliche Einführung*. Hrsg. von Elke Weick und Rainhart Lang. Wiesbaden: Gabler Verlag, S. 153–200.
- Wiseman, Alexander W., Whitty, Geoff, Tobin, Joseph und Tsui, Amy (2010). "The Uses of Evidence for Educational Policymaking: Global Contexts and International Trends". In: *Review of Research in Education* 34, S. 1–24.
- Wocken, Hans (2009). *Inklusion & Integration: Ein Versuch, die Integration vor der Abwertung und die Inklusion vor Trümereien zu bewahren*. URL: www.hans-wocken.de/Wocken-Frankfurt2009.doc.
- (2012). "Rettet die Sonderschulen? - Rettet die Menschenrechte! Ein Appell zu einem differenzierten Diskurs über Dekategorisierung". In: *Zeitschrift für Inklusion Online* 4. URL: <https://www.inklusion-online.net/index.php/inklusion-online/article/view/81/81> (besucht am 29. 03. 2018).
- Wong, Kenneth K. und Anagnostopoulos, Dorothea (1998). "Can Integrated Governance Reconstruct Teaching? Lessons learned from two low-performing Chicago High Schools". In: *Educational Policy* 12.1, S. 31–47.
- Wrase, Michael (2015). "Die Implementation des Rechts auf inklusive Schulbildung nach der UN-Behindertenrechtskonvention und ihre Evaluation aus rechtlicher Perspektive". In: *Inklusion von Schülerinnen und Schülern mit sonderpädagogischem Förderbedarf in Schulleistungserhebungen*. Hrsg. von Poldi Kuhl, Petra Stanat, Birgit Lütje-Klose, Cornelia Gresch, Hans Anand Pant und Manfred Prenzel. Wiesbaden: Springer VS, S. 41–74.
- Wrase, Michael und Piezunka, Anne (2017). *Ansprüche aus der UN-Behindertenrechtskonvention: Welche Ansprüche ergeben sich aus der UN-Behindertenrechtskonvention für Schülerinnen und Schüler mit sonderpädagogischem Förderbedarf?* URL: <https://huinklusiv.hypothesen.org/35> (besucht am 03. 07. 2017).
- Wurster, Sebastian, Feldhoff, Tobias und Gärtner, Holger (2016). "Führen verschiedene Inspektionskonzepte zu unterschiedlicher Akzeptanz und Verwendung der Ergebnisse durch Schulleitungen und Lehrkräfte?" In: *Zeitschrift für Erziehungswissenschaft*, S. 1–19.
- Wurster, Sebastian und Gärtner, Holger (2013). "Schulen im Umgang mit Schulinspektion und deren Ergebnissen". In: *Zeitschrift für Pädagogik* 59.3, S. 425–445. (Besucht am 24. 02. 2014).
- Yancy, George und Butler, Judith (2015). "What's wrong with all lives matter?" In: *New York Times*. URL: <http://opinionator.blogs.nytimes.com/2015/01/12/whats-wrong-with-all-lives-matter/>.
- Zeitler, Sigrid, Asbrand, Barbara und Heller, Nina (2013). "Steuerung durch Bildungsstandards – Bildungsstandards als Innovation zwischen Implementation und Rezeption". In: *Innovationen im Bildungswesen*. Hrsg. von Matthias Rürup und Inka Bormann. Educational Governance. Wiesbaden: VS Verlag für Sozialwissenschaften, S. 127–147.

Abbildungsverzeichnis

Abb. 3.1	Wirkungsbereiche der externen Schulevaluation	52
Abb. 5.1	Sensemaking nach Weick	76
Abb. 5.2	Sensemaking von politischen Vorgaben	83
Abb. 6.1	Schulinspektionen in Deutschland	89
Abb. 6.2	Screenshot der Auswertung	96
Abb. 8.1	Sensemaking von Inklusion	146

Tabellenverzeichnis

Tab. 6.1	Datenmaterial	88
Tab. 6.2	Antwortkategorien	92
Tab. 6.3	Kategoriensystem	95
Tab. 7.1	Operationalisierung von Leistung	113
Tab. 7.2	Definition von Inklusion in den Instrumenten	130
Tab. 7.3	Integration der Instrumente im Instrumentenset	131
Tab. 7.4	Messungen von Inklusion	135
Tab. 7.5	Deutungen der Instrumenteentwickler_innen	137
Tab. 8.1	Rahmenbedingungen beim Sensemaking	148
Tab. 8.2	Sensemaking bei Frau Elsner	183
Tab. 8.3	Sensemaking bei Herrn Pehla	185
Tab. 8.4	Sensemaking bei Herrn Plate	188
Tab. 8.5	Sensemaking bei Frau Steigleder	189

Akronyme

BW	Baden-Württemberg
BY	Bayern
BE	Berlin
BRK	Übereinkommen über die Rechte von Menschen mit Behinderung durch die Vereinten Nationen
BB	Brandenburg
HB	Bremen
BLI	Bund-Länder-Inspektion
DIMR	Deutsches Institut für Menschenrechte
EC	Economie de conventions
EFB	Elternfragebogen
HH	Hamburg
HE	Hessen
INSEE	Institut national de la statistique et des études économiques
IGLU	Internationale Grundschul-Lese-Untersuchung
IB	Inspektionsbericht
IQB	Institut zur Qualitätsentwicklung im Bildungswesen
KRK	Kinderrechtskonvention
KMK	Kultusministerkonferenz
LFB	Lehrerfragebogen
OECD	Organisation für wirtschaftliche Zusammenarbeit und Entwicklung
OFSTED	Office for Standards in Education, Children's Services and Skills
MV	Mecklenburg-Vorpommern
NI	Niedersachsen
NRW	Nordrhein-Westfalen
PISA	Programme for International Student Assessment
SL	Saarland
ST	Sachsen-Anhalt
SH	Schleswig-Holstein
SFB	Schülerfragebogen
SDG	Sustainable Development Goals
TIMSS	Trends in International Mathematics and Science Study
ÜI	Übergreifende Indikatoren
UBB	Unterrichtsbeobachtungsbogen
UN	United Nations

A Leitfaden

Der folgende Leitfaden besteht aus drei Blöcken. Im ersten Block geht es um allgemeine Fragen zur Entwicklung der Messinstrumente. Im zweiten Block geht es primär um Inklusion und die dazugehörigen Messinstrumente. Im dritten Block folgen bundesländerspezifische Fragen, die sich aus der Analyse der Messinstrumente ergeben haben. Die aufgeführte Fassung des Leitfadens entspricht der letztverwendeten Version. Während der Erhebung kam es zu einzelnen Anpassungen.

A.1 Teil 1 des Leitfadens: Allgemeine Fragen zur Entwicklung von Messinstrumenten

Themen / Bezüge zu Theorien	(Nach-)fragen
Einleitung	Danke für die Bereitschaft zum Interview Kurzvorstellung der Interviewerin Verweis auf Informationsblatt Informationen zum Aufbau des Interviews: erst allgemeine Fragen und dann Fragen zu den Instrumenten zu Inklusion
Einstiegsfrage	Können Sie sich <i>kurz vorstellen</i> und Ihre Rolle bei der Schullinspektion beschreiben?
Entstehungsprozess der Bewertungskriterien	Wie Sie wissen, interessiert mich der Entstehungsprozess der Bewertungskriterien. Können Sie mir <i>mehr über den Entstehungsprozess erzählen</i> ?
Situativer Kontext / Relevante Akteure	Von wem haben Sie in Bezug auf die <i>inhaltliche Gestaltung</i> Input erhalten? Was sind <i>weitere Akteure</i> , die ihre Meinungen einbringen können? Gab es beispielsweise <i>Vorgaben von oben</i> ? Gab es <i>Vorgaben von der Praxis</i> ? Wer hat das <i>Letztentscheidungsrecht</i> ?
Mögliche Einflussfaktoren / politische Richtlinien	Wie berücksichtigen Sie konkret die rechtlichen Vorgaben bei der Entwicklung der Bewertungskriterien?
Theorie der Wertordnungen	Welches Verständnis von guter Schule hat die Schullinspektion in ihrem Bundesland?
Theorie der Wertordnungen	Warum haben Sie sich auf dieses Verständnis geeignet? Warum ist das für Sie eine gute Schule?

Themen / Bezüge zu Theorien	(Nach-)fragen
Theorie der Wertordnungen	Was ist für Sie eine schlechte Schule?
Theorie der Wertordnungen	Inwiefern gab es bei Ihnen Diskussionen darüber, dass man unterschiedliche inhaltliche Schwerpunkte setzen müsste? Welche inhaltlichen Themen waren das? Wie kam es dann zu einer Einigung?
Theorie der Wertordnungen	Fast jedes Bundesland hat seine eigenen inhaltlichen Schwerpunkte. Wie würden sie das inhaltliche Profil von ihrem Bundesland im Vergleich zu den anderen beschreiben? Woher kommen Ihrer Meinung nach die Unterschiede?
Situativer Kontext	Welche Funktion hat die [Schulinspektion] Ihrer Meinung nach?
Bewertungssoziologische Perspektive	Was ist für Sie eine gute Evaluation? Was ist eine schlechte Evaluation?
Situativer Kontext	Spannungsfeld: Rechenschaftslegung vs. Normsetzung? Wo wird das sichtbar?
Situativer Kontext	Inwiefern können Sie die Individualität der Schule berücksichtigen? Inwiefern berücksichtigen Sie die Rahmenbedingungen der jeweiligen Schule?
Situativer Kontext	Welche Erwartungen haben Sie an die Ergebnisse der Schule?

Themen / Bezüge zu Theorien	(Nach-)fragen
Mögliche Rahmenbedingungen	Welche Ansprüche muss ein Item erfüllen, um im Unterrichtsbeobachtungsbogen aufgenommen zu werden?
Mögliche Rahmenbedingungen / Kognitive Überzeugungen und Wissen	Welche Rolle spielt z.B. die Realisierbarkeit von Bewertungskriterien? Was nehmen Sie bei der Entwicklung von Instrumenten als Herausforderung wahr? Was sind Gründe dafür? Welches Item war das?
	In Bezug auf die Vergangenheit: Wann wurde ein Standard, bzw. ein Item schon mal rausgenommen?
	Können Sie mir den Zusammenhang zwischen Orientierungsrahmen und ihren Instrumenten erklären?

A.2 Teil 2 des Leitfadens: Fragen zu Instrumenten, die dazu dienen die Umsetzung von Inklusion zu messen.

Anmoderation: Nun kommen wir zum zweiten Teil des Interviews. Hier interessiert mich das Thema Inklusion.

Thema	Frage	Nachfragen
Inklusion und Schulinspektion	Inwiefern war <i>Inklusion bei Ihnen ein Thema?</i>	
Situativer Kontext	Welche inhaltlichen Vorgaben haben Sie in Bezug auf Inklusion bekommen?	Wie sahen die Vorgaben aus? Gab es noch andere Akteure, die Ihnen inhaltliche Vorgaben gegeben haben?

Thema	Frage	Nachfragen
Eigenes Verständnis	Was versteht die Schulsinspektion unter Inklusion?	Wieso hat man sich dann auf dieses Inklusionsverständnis geeinigt?
Politische Richtlinien	Welche Rolle spielt die <i>UN-Behindertenrechtskonvention</i> für Ihre Arbeit?	Haben andere Richtlinien in Bezug auf Inklusion bei ihrer Arbeit eine Rolle gespielt?
Richtlinien	Welche Vorgaben aus dem Schulgesetz haben eine Rolle in Bezug auf Inklusion gespielt?	
Kognitive Überzeugungen und Wissen	Welche Bewertungskriterien zielen auf Inklusion ab?	
Beispiel	Ein Merkmal von inklusiver Schule kann sein, dass alle Schülerinnen und Schüler individuelle Förderpläne haben. Dieses Merkmal taucht meines Wissens bei den Bewertungskriterien vieler Bundesländer nicht auf. Was könnten Ihrer Meinung nach Gründe dafür sein?	Warum taucht es bei Ihnen nicht auf?
Bezug zu Teil 1 Situativer Kontext/ Politische Richtlinien	Vorhin haben Sie erwähnt, dass XYZ bei der Entwicklung von Bewertungskriterien wichtig ist – inwiefern trifft dies auch auf das Thema Inklusion zu?	
Theorie der Werteordnung	Was glauben Sie... Könnte eine Schule, <i>die NICHT inklusiv ist</i> , gut bei Ihnen abschneiden?	

Thema	Frage	Nachfragen
Abschluss	<i>Gibt es noch etwas von Ihrer Seite, dass Sie gerne ergänzen würden? Danke für die Teilnahme</i>	

A.3 Teil 3: Bundesländerspezifische Fragen

Vor der Durchführung der Interviews hat die Autorin die Messinstrumente analysiert und die Verfahren in den einzelnen Bundesländern näher betrachtet. In diesem Zusammenhang sind Nachfragen entstanden, die in diesem Teil des Interviews gestellt wurden. Aus Datenschutzgründen dürfen die bundeslandspezifischen Nachfragen nicht aufgeführt werden.

B Transkriptionsregeln

- Wortwörtliche Transkription: z.B. „haste“ anstatt „hast du“
- Festhalten, wer spricht (Verwendung von Abkürzungen möglich):
 - AP = Interviewerin
 - IP = Interviewte Person
- Bei der Verschriftlichung Groß- und Kleinschreibung verwenden
- Pausen werden transkribiert
 - (.) = Mikropause (< 1sec)
 - (1) = Pausen in Sekundenlänge
 - - = Wort oder Satzabbruch
- Nichtverbale Äußerungen (z.B. Lachen, Räuspern, Husten, Stottern) werden transkribiert
 - <lacht>
 - <hustet>
 - Mhm, hmhm
- Namen und Orte werden anonymisiert, z.B.
 - [Name des Interviewten]
 - [Name der Hochschule]
- Tonhöhenbewegungen werden nicht transkribiert
- Zahlen werden als Ziffern notiert, außer es gibt lange Pausen oder nonverbale Äußerungen dazwischen, z.B. Drei ähm und zwanzig
- Aussagen in Englisch werden in englischer Sprache notiert und die englische Rechtschreibung wird berücksichtigt.
- Falls Aussagen nicht deutlich erkennbar sind, wird dies gekennzeichnet, z.B. so: [nicht verstanden in Transkription]

C Danksagungen

Bei der Promotion haben mich viele Menschen unterstützt, denen ich auf diesem Wege gerne danken möchte:

Als erstes möchte ich meine Doktormutter Vera Moser nennen: Herzlichen Dank für deine intensive und fürsorgliche Betreuung – Ich konnte dich in den letzten Jahren jederzeit um Rat fragen und bin dir für deine Unterstützung bei der Dissertation und allem, was drumherum passiert ist, sehr dankbar. Ich möchte dir auch für dein Vertrauen und dem damit einhergehenden Freiraum danken! Zugleich aber auch für die ermutigenden Impulse zum Weiterarbeiten. Durch die Einbindung in das Graduiertenkolleg und in das Zentrum für Inklusionsforschung Berlin (ZfIB) haben sich für mich sehr viele Chancen ergeben. Ein großes Dankeschön geht auch an meine Zweitbetreuerin Melanie Fabel-Lamla, insbesondere für Ihr aufmerksames Zuhören bei unseren Treffen und hilfreichen Nachfragen, die häufig noch lange nachwirkten. Danke, dass auch Sie sich viel Zeit für meine Arbeit genommen haben und danke für Ihre Einschätzungen, was ich innerhalb einer Dissertation leisten sollte und was wiederum nicht machbar ist.

Ein großer Dank geht an das Wissenschaftszentrum Berlin für Sozialforschung (WZB) bzw. an all die Personen, mit denen ich dort in den letzten Jahren zusammengearbeitet habe. Nadine Bernhard, Justin Powell, Heike Solga und Lukas Graf möchte ich dafür danken, dass ich durch sie im Rahmen des INVEST-Projekts einen ersten Einblick in die empirische Bildungsforschung erhalten habe. Ein besonderer Dank geht an Reinhard Pollak, der mich stets darin bestärkt hat neue Herausforderungen anzunehmen und zugleich mir die Möglichkeit gegeben hat diese zu realisieren, wie den Aufenthalt an der Northwestern University in Evanston/Chicago. Danke für deine Offenheit gegenüber meinen Projektideen sowie deine Unterstützung und Hinweise, wenn ich in Bezug auf die Dissertation, meine finanzielle Absicherung oder andere Forschungsprojekte Hilfe benötigte. In diesem Zusammenhang möchte ich generell dem WZB für die Chancen danken, die ich erhalten habe, z.B. der Unterstützung bei Publikationen, Weiterbildungen, Übergangsförderungen, Austausch mit Politiker_innen usw.

Diese Arbeit wäre nicht möglich gewesen ohne die Unterstützung der Mitarbeiter_innen der Schulinspektionen: Herzlichen Dank für die Bereitschaft Interviews zu geben und Materialien bereit zu stellen. Ich habe insbesondere die Teilnahme an den jährlichen Treffen als großen Gewinn für meine Arbeit wahrgenommen und habe mich dabei sehr über Ihr Interesse an meiner Arbeit gefreut. Es war toll in Soest, Wiesbaden und Speyer dabei sein zu dürfen. In diesem Zusammenhang auch Danke an Fanny Oehme, die mir nicht nur einen vereinfachten Kontakt zum Feld ermöglicht hat, sondern als Freundin und Expertin zu Schulinspektionen diese Arbeit über den gesamten Zeitraum intensiv begleitet hat.

Der Stiftung Universität Hildesheim möchte ich dafür danken, dass sie mir durch das dreijährige Stipendium des Landes Niedersachsen die Möglichkeit gegeben hat, meine eigene Fragestellung zu entwickeln und diese selbstständig zu verfolgen. Dies habe ich als großes Privileg wahrgenommen. Danke auch dem Zentrum für Bildungsintegration und insbesondere Viola Georgi für den Freiraum, der uns Stipendiat_innen gegeben wurde, und die anregenden Veranstaltungen, die in Hildesheim realisiert wurden. Das Zentrum für Bildungsintegration hat mir ermöglicht mit Forscher_innen aus verschiedenen Disziplinen und unterschiedlichen Ländern in den Dialog zu treten und dabei bestehende Annahmen kritisch zu hinterfragen. In diesem Zusammenhang

war auch der Austausch mit den anderen Kollegiat_innen (insbesondere Sabine Hoffmann, Dilek Ikiz, Seyran Bostancı und Nicholas Stone) sehr gewinnbringend.

Des Weiteren hatte ich neben dem Promotionskolleg in Hildesheim auch als assoziiertes Mitglied im Kolleg „Inklusion-Bildung-Schule“ an der Humboldt Universität zu Berlin die Möglichkeit regelmäßig meine Zwischenstände zu präsentieren und zu diskutieren. Ich möchte den anderen Kollegiat_innen (Anja Akhoondi, Angelika Bengel, Laura Capellmann, Sophie Görtler, Jona Garz, Florian Schrumpf, Nele Rordorf, Dana Tegge, Robert Aust, Aylin Jordan, Thea Nesyba, Laura Röbenack, Maik Walm, Julia Wiebigke und Christian Brüggemann) sowie der Leitungsrunde (Vera Moser, Detlef Pech, Rita Nikolai, Ada Sasse und Anna Moldenhauer) für ihre wertvollen Hinweise danken. Es hat Spaß gemacht mit den Mitgliedern des Kollegs unterwegs zu sein, z.B. bei den Inklusionsforschertagungen in Halle, Gießen und Linz, der ECER-Tagung in Kopenhagen, der DGfE-Tagung in Essen oder auf Schreibwoche in Brandenburg. In diesem Zusammenhang auch Danke an Marina Egger, die unseren leicht chaotischen Haufen stets zusammenhält. Des Weiteren möchte ich auch Annedore Prengel danken, die mich sowohl in Bezug auf die Dissertation als auch persönlich sehr geprägt hat. Es macht unheimlich viel Spaß mit dir zusammenzuarbeiten und ich lerne sehr viel. Danke, dass es dich gibt.

Herzlichen Dank für ihre inhaltliche Expertise und kritische Rückmeldungen

- zu Inklusion an Annedore Prengel, Elisabeth Plate, Justin Powell, Michael Wrase, Michael Grosche, Cornelia Gresch und Sven Degenhardt,
- zu Schulinspektionen: Fanny Oehme, Moritz Sowada, Maike Lambrecht und Fabian Dietrich,
- zum Sensemaking-Ansatz: James Spillane, Jeannette Colyvas sowie Cynthia Coburn,
- zu bewertungssoziologischen Fragestellungen: Wendy Espeland, Simone Bloem, Anne Krüger und Stefan Beljean,
- zu qualitativen Methoden: Nadine Bernhard sowie der qualitativen Forschungswerkstatt am WZB.

Danke für Korrekturen, Unterstützung bei den Transkriptionen, Recherchen, Feedback, Leidensgenossenschaften und/oder emotionale Unterstützung: Agnieszka Althaber, Ariane Berthoin-Antal, Jonna Blanck, Yvonne Blos, Kristin Bothur, Sophie Drebes, Cindy Fitzner, Roberta Gärtig, Julia Gasterstädt, Hans Gerhardt, Sophie Görtler, Tine Hanrieder, Susanne Hänsch, Silvia Höhne, Johanna Hofmann, Vera Klar-Winter, Julia Knopp, Ralf Künster, Franziska Lui, Clara Overweg, Josefine Matysiak, Laura Menze, Thea Nesyba, Juliane Pehla, Sylvia Pichorner, Bärbel Pietraß, Sigrid Piezunka, Katja Pomianowicz, Hubertus Redlich, Tom Reichelt, Christian Roth, Michaela Roth, Lea Schäfer, Tina Schaffus, Angela Schrauber, Benjamin Schulz, Sören Stapel, Jennifer Weitz, Ellen von den Driesch, Katharina von der Ohe, Sebastian Wendel, Katja Zehbe, Julia Zeilinger, Katji Zauner, dem AAM-Neps-Colloquium des WZB und dem Berliner Colloquium für Inklusion.

Zum Schluss möchte ich meiner Familie danken, der ich diese Arbeit gerne widmen möchte. Danke an meine Mutter Hille Piezunka, meinen Vater Hartmut Piezunka und meinen Bruder Henning für ihre unglaubliche Unterstützung in all den Jahren, dass sie in schwierigen Momenten immer für mich da sind, für ihr Hinterfragen von Dingen und ihre Neugier sowie all die anderen Werte, die sie mir vorgelebt haben bzw. weiterhin vorleben. Ich bin traurig darüber, dass ich nicht mehr die Möglichkeit habe diese Arbeit mit meinem Vater, der 2008 verstorben ist, zu diskutieren und darüber zu streiten.

Ein besonderer Dank geht an meinen Mann Christian für seine unendliche Geduld mit mir und insbesondere für seine Unterstützung, bei all den kleinen und großen Krisen, die sich

während der Dissertation ereignet haben. Danke fürs Zuhören, in dem Arm nehmen und Mitpacken! Mit großer Dankbarkeit und Glück denke ich daran zurück, was wir in den letzten Jahren gemeinsam erlebt haben, insbesondere in Chicago und an der Loire. Ich bin für deine große Unterstützung bei der Dissertationsarbeit und all dem, was zeitgleich auch noch passiert ist, sehr dankbar.

Wie misst man die Qualität von Schule und was versteht man überhaupt unter einer „guten“ Schule? Mit diesen Fragen setzen sich im Fall von Deutschland die bundeslandspezifischen Schulinspektionen auseinander. Sie evaluieren mithilfe von Befragungen und standardisierten Beobachtungen regelmäßig die Qualität von Einzelschulen. Im Rahmen der Arbeit wird untersucht, inwiefern Schulinspektionen unter einer guten Schule eine inklusive Schule verstehen und wie die dazugehörigen Messinstrumente entstanden sind. Da es an einer einheitlichen Definition von schulischer Inklusion fehlt, wird im Rahmen der Arbeit auch berücksichtigt, was die Schulinspektionen in den einzelnen Bundesländern darunter verstehen und wie sie die gesetzliche Vorgabe in ihren Messinstrumenten interpretiert haben. Hierbei wird auf Arbeiten aus der Inklusionsforschung (z.B. Prengel; Hinz) sowie auf Arbeiten aus der Bewertungssoziologie (z.B. Boltanski & Thévenot) zurückgegriffen. Des Weiteren wird mithilfe des organisationssoziologischen Sensemaking-Ansatzes (Weick; Spillane) analysiert, welche Rahmenbedingungen bei der Entwicklung der Messinstrumente eine Rolle spielten, z.B. bildungspolitische Konstellationen sowie persönliche Erfahrungen.

Diese Dissertation wurde 2019 mit dem Wissenschaftspreis der Sektion Sonderpädagogik der DGfE ausgezeichnet.

Interdisziplinäre Beiträge zur Inklusionsforschung

Die Herausgeberin und der Herausgeber



Anne Piezunka, Jahrgang 1986, ist wissenschaftliche Mitarbeiterin am Wissenschaftszentrum Berlin für Sozialforschung und an der Akademie für bildungsgeschichtliche und zeitdiagnostische Forschung an der Universität Potsdam. Zu ihren Forschungsinteressen gehören: Umsetzung von Inklusion, Evaluationsprozesse, Educational Governance.

