Measuring Exposure and Attention to Media and Communication

Solutions to Wicked Problems

Peter Neijens
Theo Araujo
Judith Möller
Claes H. de Vreese

Amsterdam
University
Press

Measuring Exposure and Attention to Media and Communication

What Experts Say about this Book

Michael Slater, Social and Behavioral Sciences Distinguished Professor, School of Communication, The Ohio State University, USA

"Assessment of exposure to messages is fundamental to the study of mediated communication, and there are few variables concerning social experience that are trickier to operationalize, especially given our rapidly evolving communication environment. This book concisely summarizes a wide range of the state-of-the-art in approaches to assessing media exposure, ranging from ratings to eye tracking, from self-report to digital traces, from ecological momentary assessment to media buy impressions. Accordingly, the book addresses many of the complexities associated with accessing media content through apps, social media, and other software-driven interfaces. The reference list is impressively comprehensive and useful in pointing the way to more in-depth explorations of specific measurement methods. This book will be an invaluable resource for training graduate students and for exploring research design alternatives by faculty and industry researchers in communication, media psychology, and allied fields."

Veronika Karnowski, Chair of Media Communication, Chemnitz University of Technology, Germany

"Media exposure is one of the — if not even THE — core concept of our field. Hence, we need to talk about its conceptualization and measurement. This book is the synthesis of many years of research and discussions about this concept, not only but mainly at ASCoR. In this book, Peter Neijens, Theo Araujo, Judith Möller, and Claes de Vreese gather the state of the art of conceptualizing and measuring media exposure. From self-report measures over digital trace data, observations, eye tracking, and neurobiological measures to ecological measures, they cover pretty much everything our fields' methodological toolkit currently offers to tackle media exposure. And they give insightful recommendations on furthering our debates on and measurement of media exposure. I am confident that this book will provide an excellent resource for many, from students to experienced experts in the field, and that it will instill much-needed future discussions and research on the conceptualization and measurement of this core construct."

Measuring Exposure and Attention to Media and Communication

Solutions to Wicked Problems

Peter Neijens
Theo Araujo
Judith Möller
Claes H. de Vreese

With contributions from
Claire M. Segijn & Emily Vraga, and
Frederic R. Hopp & Bert N. Bakker

Amsterdam University Press

The publication of this book is made possible by a grant from the Research Priority Area (RPA) "Communication", funded by the Faculty of Social and Behavioural Sciences (University of Amsterdam). Claes de Vreese and Peter Neijens were responsible for the Media Exposure project within this RPA. From 2018, this project was part of the Digital Communication Methods Lab (Digicomlab) led by Claes de Vreese, Theo Araujo and Judith Möller. Claire M. Segijn and Emily Vraga contributed to this book because of their specialist knowledge of eye tracking; Frederic R. Hopp and Bert N. Bakker because of their specialist knowledge of neurobiological measures.

Cover design: Coördesign, Leiden
Lay-out: Crius Group, Hulshout

ISBN 978 94 6372 317 6
e-ISBN 978 90 4855 546 8
DOI 10.5117/9789463723176
NUR 670

Table of Contents

Acknowledgments 7

Preface 9

1. Introduction 11

2. Conceptualizing Media Contact 15

3. Quality Criteria for Media Exposure Measures 33

4. Self-Report Measures 41

5. Digital Trace Data 87

6. Observation 105

7. Eye Tracking 113
Claire M. Segijn and Emily Vraga

8. Neurobiological Measures 129
Frederic R. Hopp and Bert N. Bakker

9. Ecological Measures 145

10. Recommendations 151

Index 161

Acknowledgments

Thank you:
- All scholars who have studied media exposure and its measurement. Their studies form the basis of this book.
- University of Amsterdam, Faculty of Social and Behavioural Sciences, Research Priority Area Communication, and the Amsterdam School of Communication Research for financial support to our project.
- Edith Smit (The Amsterdam School of Communication Research, University of Amsterdam) and Anita Elberse (Harvard Business School, USA) for your contributions in the early stages of the project.
- Student-assistants Lori Meester and Anne Brons.
- Yuzhi Lu for performing the writing checks.
- All participants in our expert meetings (see Preface).

Preface

The conceptualization, operationalization, and measurement of media exposure have had our attention for quite some time. The reason is simple: valid and reliable measurement of media and communication exposure is crucial for communication science, psychology, political science, sociology, pedagogy, economics, and law, and the practitioners in media, communication, and information. At the same time this is a "wicked problem" for which there are no simple solutions.

Over the years, we have developed several activities to improve the quality of media exposure measurement. These activities under the umbrella of the Research Priority Area "Communication" of The Amsterdam School of Communication Research (ASCoR) and the University of Amsterdam included a variety of activities. We repeatedly carried out studies and overviews in which existing, and newly developed measures were put to the test (see the References in each chapter). We have also developed a website with media exposure measures and their characteristics that allowed researchers to search for extant media exposure measures, obtain available information about the quality and application of these measures, and add new measures and information to this overview resource. It was meant as a simple, practical, and shared tool assisting researchers to assess what kinds of measures are most suited to their research interests. Considering the dynamic pace of changes in media exposure measurement, we decided to discontinue the maintenance, but it has long served as a useful global resource.

Furthermore, we organized several expert meetings (in 2014, 2016, and 2022) where international experts shared and discussed their research on this topic. Participants included Ava Francesca Battocchio (Michigan State, USA), Jeffrey Boase (University of Toronto, Canada), Susanna Dilliplane (Aspen Institute, USA), Karin Fikkers (Utrecht University), Veronika Karnowski (LMU Munich, Germany), Frans Kok (Nationaal Media Onderzoek / Frans Kok Projects, the Netherlands), Michael LaCour (UCLA, USA), Ed Malthouse (Northwestern University, USA), Ericka Menchen Trevino (EMT LLC, USA), Douglas A. Parry (Stellenbosch University, South Africa), Elizabeth Saad Correa (University of São Paulo, Brazil), Dhavan Shah (University of Wisconsin–Madison, USA), Nathalie Sonck (Kantar, the Netherlands), Talia Stroud (University of Texas at Austin, USA), Harsh Taneja (University of Illinois, USA), and Kjerstin Thorson (Michigan State, USA). We thank them for their invaluable contributions and the great discussions. We also thank the participants in the Annual Conference of the Media Audiences and Effects Division of the German Communication Association (DGPuK) in 2016, and the participants in the Key Concepts workshop organized by Yariv Tsfati in 2023 for the reflections on current challenges to media exposure measurements.

We edited a special issue of *Communication Methods and Measures* (published in 2016) with contributions from Erik Albæk (University of Southern Denmark), Kim Andersen (University of Southern Denmark), Jason Barabas (Stony Brook University, USA), Susan Banducci (University of Exeter, UK), Jeffrey Boase (University of Toronto, Canada), Leticia Bode (Georgetown University, USA), Robert Hornik (Annenberg School, USA), Jennifer Jerit (Stony Brook University, USA), Jiaying Liu (Annenberg School, USA), Jeff Niederdeppe (Cornell University, USA), Jakob Ohme (University of Southern Denmark), William Pollock (Stony Brook University, USA), Martijn Schoonvelde (Vrije Universiteit Amsterdam, the Netherlands), Michael Slater (Ohio State, USA), Daniel Stevens (University of Exeter, UK), Harsh Taneja (University of Missouri, USA), Sonya Troller-Renfree (University of Maryland, USA), Emily Vraga (George Mason University, USA), and James Webster (Northwestern, USA). This special issue is now an often-cited resource in this space.

The current book feels like a "full circle" project for us. The aim is sharing our knowledge gathered in these projects and from the literature. The book provides an overview of the different measurement methods, from the good old self-reports to the recent data donation methods. We discuss their pros and cons and empirical performance. We believe that such an overview is useful, as knowledge and understanding of these methods and issues has been spread over many publications, academic and applied, from different disciplinary fields, over many years. Our objective is to highlight all important themes and issues. Given the large number of studies, it is impossible to list each publication individually, but we have tried to do justice to previous authors as much as possible.

Studies and scholarly discussions on the best way of measuring media exposure are fascinating and inspiring and we hope that the book passes this spirit on to (new) colleagues and students in the field. Inspiration of many is much needed for the further improvement of these methods, as we realize that developments in technology will not only create new measurement problems, but also opportunities to solve them.

1. Introduction

Picture your day and think about your encounters with media and communication. You probably start your day checking your mobile phone for news and messages from your friends, you listen to streaming music while you exercise, you read the news on your tablet during breakfast, you are confronted with advertisements while you listen to podcasts or radio streaming on your way to work, you hear fellow travelers discussing (news) in breaks from their cell phones, you stream music in your office while checking messages from the team manager and colleagues on your laptop, and you play a game on your mobile phone in the lunch break. In the evening you watch a soccer game with friends in a bar, while passing a personalized public space ad that interacts with your smart watch, you search for information about a new mobile phone, browse through pictures shared on social media, and maybe you watch a bit of television in between a Netflix show, and not to forget: all this in combination with regularly checking your phone for news and messages from your friends. All these activities are worth studying on their own. But they also have a known impact on many different variables on the individual and the societal level: ranging from mental and physical well-being to the polarization of the public sphere.

Media and communication exposure (MCE) — that is "the extent to which audience members have encountered specific messages or classes of messages/media content" (Slater, 2004, p. 168) — is a crucial concept in studies on media audiences and effects in communication science, psychology, political science, sociology, pedagogy, law, and economics. This concept has different roles in these studies: it is a *dependent* variable in theories and studies on media use; a *mediator* in selective exposure theories that specify that persons with certain characteristics seek out specific media that subsequently impact them; an *independent* variable in media effects theories, and a *moderator* in theories suggesting that exposure interacts with individual level and contextual factors.

Media exposure concepts are not only highly relevant for scientific studies, but also for the media industry. Aggregated media exposure data such as circulation, ratings and reach are indispensable for media programming

Neijens, P., Araujo, T., Möller, J., & de Vreese, C. H. (2024). Introduction. In P. Neijens, T. Araujo, J. Möller, & C. H. de Vreese, *Measuring exposure and attention to media and communication. Solutions to wicked problems* (pp. 11-14). Amsterdam University Press. https://doi.org/10.5117/9789463723176_ch01

and publishing decisions, and as a currency for buying and selling of media space for political, commercial, and health campaigns. Media exposure data such as exposure to "questionable media content" by "vulnerable groups" are important for policy decisions and parenting advice.

Media exposure is thus a crucial concept. But it is also complex as it relates to a wide variety of behaviors, platforms, devices, content, and situations. Especially in today's media landscape which is characterized by an abundant number and variety of traditional and new media (offline, online, mobile, meta), communicators (professional newsmakers such as journalists, PR professionals, advertisers, as well as non-professional "senders" such as other media users, social media influencers, social contacts, and algorithms), (blurred) content categories (ranges from news, information, political comedy, advertising, health information, and entertainment to personal experiences), and modalities (text, images, voice, video), increasingly "on demand," interactive, and "unbundled" in the form of single articles, music tracks, and videos distributed through different channels. Media consumption varies from active, deep, and focused to passive, habitual, ephemeral, unconscious, short and superficial; is highly personalized and characterized by media multitasking, time shifting, less stable repertoires; and is consumed on different devices in different situations at home, at work, and while traveling.

These multi-faceted aspects have led to the felt necessity to apply different conceptualizations and operationalizations of "exposure." Also, exposure measurement may take many different forms. They can be clustered into categories: self-report measures (e.g., questionnaires, diaries, think-aloud) which are based on media users' own assessments (Chapter 4), digital trace data which is registered by a device (e.g., meters, mobile apps) (Chapter 5), (human) observation which requires observers (Chapter 6), eye tracking (Chapter 7), neurobiological measures including psychophysiological (e.g., heart rate, skin conductance) and fMRI data (Chapter 8), and ecological measures which are derived from statistical data (Chapter 9).

Measurement of media exposure was already difficult in the media landscape of 25 years ago. But today it is even harder. Discussions about the conceptualization and measurement of media exposure and methodological studies into the quality of the measures have a long history, reaching an all-time high in the past decade. This trend reflects the explosively increased amount and variety of media products and modes or contexts of consumption, the associated theoretical, political, and social interest in these phenomena, and the increased possibilities of (mobile) tracking methods brought by digitalization.

It is difficult to make informed choices on the measurement of MCE in view of the many possibilities and the scattered knowledge on their advantages and disadvantages. This book therefore attempts to integrate the insights by providing an overview of:

– the ways in which MCE can be *conceptualized* and *operationalized* (Chapter 2),
– the ways in which MCE can be *measured*, their pros and cons, and performance (Chapter 3–9),
– and *recommendations* for the application and further development of these methods (Chapter 10).

The chapters on the different exposure method differ in length, which is related to differences in complexity, popularity, and numbers of methodological studies into the method. The book also includes an extensive bibliography — with references to in-depth insights into specific aspects of media exposure measurement.

We are well aware that our book will neither be the ultimate reference work on the subject nor settle longstanding disputes over different approaches or shortcomings of specific measures. We wouldn't have the audacity to let this be the ambition. Instead, we hope that the book will summarize and consolidate what we know, provide guidance on what considerations are important, and provide a framework for evaluating specific measures and broader developments.

Reference

Slater, M. D. (2004). Operationalizing and analyzing exposure: The foundation of media effects research. *Journalism & Mass Communication Quarterly, 81*(1), 168–183. https://doi.org/10.1177/107769900408100112

About the authors

Peter Neijens is an Honorary Fellow and Emeritus Professor at ASCoR. He held the Chair in Media and Persuasion at the University of Amsterdam until 2019. His research focuses on persuasive communication, in particular media use and effects in advertising and public opinion.

Theo Araujo is a Full Professor of Media, Organisations and Society, and Scientific Director of ASCoR. His research investigates the dynamic interplay

between media and organizations, and what it means for society, with a special focus on trust and technology. His research has also a methodological focus on computational communication science and the implementation of large-scale data collection and analysis for communication research.

Judith Möller holds the chair of "Empirical Communication Research, especially Media Use and Social Media Effects" at the University of Hamburg and the Leibniz Institute for Media Research | Hans Bredow Institute (HBI). Before 2023 she was an Associate Professor of Political Communication at the Amsterdam School of Communication Research.

Claes H. de Vreese is a Distinguished University Professor of AI and Society with a special emphasis on media and democracy. He co-directs the Gravitation Program AlgoSoc and the AI, Media and Democracy Lab. He holds the chair in Political Communication at the Amsterdam School of Communication Research.

2. Conceptualizing Media Contact

Abstract
Conceptual issues lie at the heart of the debate on media and communication exposure. We discuss these in this chapter. In particular, we pay attention to (1) the different concepts related to media contact, (2) medium and communication levels, (3) institutionalized media audiences, and (4) metrics.

Keywords: exposure, attention, reception, medium and communication levels, institutionalized media audiences, metrics

2.1 Media contact concepts

Media *exposure* is defined as proximity to a message (Lee, 2009), or “the extent to which audience members have encountered specific messages or classes of messages/media content” (Slater, 2004, p. 168), or any situation in which a person comes into contact with particular events or news stories through one or another medium (Allen & Waks, 1986). Information processing approaches take a broader view of media exposure. These approaches view media exposure as a dynamic process that includes psychophysiological, cognitive, and affective responses. These responses are considered relevant because they can influence the effects of people’s exposure to media (Nabi & Oliver, 2009; Oliver et al., 2020; Valkenburg & Peter, 2013a). For example, in the field of advertising “context-induced psychological responses, such as involvement elicited by a documentary, happiness caused by a sitcom, or sadness generated by a drama series, are considered to have an important impact on advertising processing” (Moorman et al., 2005, p. 49). In this section we give a brief overview of the concepts related to the various physiological and mental processes that occur during media exposure. We focus on attention, involvement, engagement, valence, arousal, emotional responses, and experiences.

Neijens, P., Araujo, T., Möller, J., & de Vreese, C. H. (2024). Conceptualizing media contact. In P. Neijens, T. Araujo, J. Möller, & C. H. de Vreese, *Measuring exposure and attention to media and communication. Solutions to wicked problems* (pp. 15-32). Amsterdam University Press. https://doi.org/10.5117/9789463723176_ch02

Attention

Associated with different attention levels are questions about operationalizations that are used to describe the media and communication "encounter", such as Opportunity to See (OTS), open eyes/ears in front of media content, presence in the room, looking, seeing, watching, listening, hearing, reading, or using. For instance, what behavior actually counts as "watching"? Is it about being present in the room (without eyes on screen), facing a screen, looking at the screen, or watching a whole program, or episode, with attention? Similar questions play a role in readership research that focus on print media: what behavior actually counts as reading? Is skimming reading? Does reading need to be done with full attention? And how long does a reader need to perform this activity to be considered to have read a newspaper?

User's *attention* to media and communication is defined in different ways, with components as alertness, selectivity, and processing capacity (Moray, 1970; Posner & Boies, 1971). Attention definitions include "the ability to focus on [the message] and also to suppress attention to other things in the environment" (Bellman et al., 2019, p. 295), "psychological involvement" (Webster & Wakshlag, 1985), or "mental effort" (Chaffee & Schleuder, 1986).

Although a relevant variable for media exposure research, the role of attention in media effect models should be interpreted with caution, because attention "is inherently confounded with the prior knowledge, interest, and attitudes that give rise to attention. Such prior variables might independently be related to the outcome of interest independent of exposure" (Slater, 2004, p. 174, see also Romantan et al., 2008).

Greenwald and Leavitt (1984) distinguish four levels of message processing referring to different attention levels: preattention (scanning media content in a subconscious and automatic way, filtering all incoming information), focal attention (where enough attention is paid to determine what the content is about), comprehension (the message is analyzed to assign meaning), and elaboration (connection to existing knowledge structures, generating personal connections and imagery) (Finn, 1988; Shapiro et al., 1997; Smit et al., 2013).[1] Eye tracking (see Chapter 7) and psychophysiological methods (see Chapter 8) are usually considered appropriate methods for lower levels

1 It is generally believed that two processes drive attention: a bottom-up, stimulus driven process in which features of media content, such as size and shape, almost automatically capture attention, even when someone is not actively searching for them, and a top-down, conscious process in which person-related factors (e.g., interest, motivation) encourage people to enhance attention (Geiger & Newhagen, 1993; Pieters & Wedel, 2004, 2007; Rosbergen et al., 1997). See also Smit et al., 2013.

of attention, self-reports including thought listing (Petty et al., 1983; see also Chapter 4) for (higher) attention levels that produce at least a minimal memory trace.

The relationship between attention and exposure is discussed by W. Potter (2008): "The idea of exposure addresses whether a person was physically exposed to a message or not, while the idea of attention addresses how much cognitive effort of concentration the person devoted to the message. Scholars appear to use this basic idea of attention as a measure of degree of exposure. As people move higher on the attention continuum, they expend more effort at perceiving and processing messages; they increase their awareness of that message; and that increased awareness along with increased cognitive effort makes it more likely that the message will be encoded into memory and more easily retrieved later" (p. 153).

W. Potter proposes an alternative model in which attention is not a single continuum but a variety of alternative states that differ from one another in kind rather than by degree. He differentiates an automatic state (in which the media user unconsciously processes messages), a transported state (in which the user is "lost" in a story, receives messages, compare the concept of "flow" (Csikszentmihalyi, 1988), an attentive state (in which an individual consciously interacts with a message), and a self-reflexive state (in which an individual consciously processes a message and reflects on that process). These four qualitatively different exposure states have implications for research design and measurement, says W. Potter: "This consideration of exposure states can help researchers make a more reasoned assessment of the types of knowledge their potential respondents do not have. Any measure, no matter how well crafted, that attempts to tap into knowledge that the respondents do not have will never achieve validity. Instead, researchers need to focus on the types of information that their respondents are likely to have, then attempt to design measures that access that information as cleanly as possible" (p. 164). He advises researchers to create conditions in their experiments and surveys in which the exposure status of participants is comparable to that in their daily life (see also Fikkers et al., 2015; Valkenburg & Peter, 2013a).

2.2 Involvement and engagement

Involvement is defined by Zaichkowsky (1985) as "a person's perceived relevance of the object based on inherent needs, values, and interests" (p. 342). Based on their meta-analysis, Johnson and Eagly (1989) characterize involvement as "a motivational state induced by an association between an activated attitude and the self-concept" (p. 290), and distinguish value relevant, impression relevant, or outcome relevant involvement.

Related to involvement is *engagement* (Advertising Research Foundation, 2006a; Araujo, et al., 2020; Napoli, 2012; Webster et al., 2014), "a state of being involved, occupied, fully absorbed, or engrossed in something" (Higgins & Scholer, 2009, p. 102), or "a multilevel, multidimensional construct that emerges from the thoughts and feelings about one or more rich experiences involved in reaching a personal goal" (Calder et al., 2016, p. 40).

Arousal, valence, and emotional responses

Arousal (intensity) and valence (positive or negative) are two core dimensions of immediate affective media responses (Russell, 1980; see also Chapter 8). Examples of emotional responses are feelings of anger, happiness, and sadness (Nabi, 2009).

Media experience

Media experience is defined by Bronner & Neijens (2006, p. 83) as "emotional, intuitive perceptions that people have while using the media" (after Koppe, 1998). The concept captures "the qualitative thoughts and feelings people have about a [medium] — what it means to like and use [the medium] from their perspective" (Malthouse et al., 2007, p. 8; see also Malthouse et al., 2003). Van der Wal et al. (2022, p. 20) — following Grady et al. (2022, p. 525) — define: "a media experience encompasses both media exposure and its attendant affective and cognitive processes, such as appraisal."

Related concepts

- Media *use* — important in Uses and Gratifications (e.g., Katz et al., 1973) — is defined by McQuail (1994, p. 303) as "the act of choosing or attending to media," where "choosing" refers to audience activity before exposure ("selectivity"), and "attending" refers to audience activity during exposure ("involvement").
- Media *reception* indicates "getting" the message (Lee, 2009), as it "requires attending to, comprehending, and retaining [remembering] news" (Price & Zaller, 1993, p. 134).

- *Generating content* includes contributing to media content (e.g., rating, commenting) and creating media content (actively producing and publishing content, e.g., writing reviews) (Muntinga et al., 2011).
- Media *effects* include structural changes or reinforcements in cognitions, emotions, attitudes, beliefs, physiology, or behavior (W. Potter, 2011, 2012). Media effect variables are beyond the scope of this book.

2.3 Medium and communication levels

In addition to issues related to "encountered" that we discussed above, there is another aspect in Slater's definition that needs further consideration: "specific messages or classes of messages/media content." Theories and studies on the effects of media and communication can refer to different levels of media content, for instance, "medium type" (e.g., television, newspapers, radio), a specific "vehicle" (e.g., *The New York Times*, *BBC News at Ten*, *Newsweek*), a specific "issue" of a newspaper, magazine, program, a specific "unit" (front page, back page), a specific article or advertisement (Advertising Research Foundation, 1961; Harvey, 1997), a specific "genre" (e.g., news, soap operas, ads), or a specific topic ("violence").

The content a person is exposed to can be further detailed. For example, Fikkers et al. (2015) argue that researchers should go a step further than asking for exposure to media violence and develop measures that are sensitive to the different types of violence that are theoretically relevant. Even more detail can be reached by associating the findings of a content analysis of the content to which a person is exposed with measures of exposure, the so-called "linkage" method (see Chapter 4).

In addition to media content, the user's media encounter can be characterized by channel and interaction indicators that are central in the hierarchical taxonomy of computer-mediated communication (CMC) of Meier and Reinecke (2021). Table 2.1 shows that the authors distinguish channel levels such as device, application type, brand and features of the app, and communication levels such as type of interaction and type of message.

In this context, van der Wal et al. (2022) adopt an activity-based framework of social media activities — suggested by Weinstein (2018) and Yang et al. (2021) — that "revolve around relational interactions (direct messaging), self-presentation/expression (posting and broadcasting), and browsing" (p. 21).

Table 2.1 The hierarchical CMC taxonomy

Level of analysis	*Examples*
Channel centered	
Device	smartphone, tablet
Type of application	SNS, social media, email
Branded app	Facebook, Instagram
Features	status update, profile, messenger, chat
Communication centered	
Interaction	one-to-one, one-to-many, network size, ties, active, passive
Message	mode (text, voice), content (valence, topic)

Source: Meier and Reinecke (2021, p. 1185).

Exposure concepts can be measured at each of these media and communication levels, and it is important to take these levels into account when designing studies, examining theories, or comparing research findings. To illustrate this: the often-used indicator “screen time” may hide a diversity of content and interactions, such as browsing through live streams, posting videos, playing games, selling something, or creating a poll (Griffioen et al., 2020b; Parry et al., 2021).

2.4 The media industry: institutionalized media audience

Data for the media industry is negotiated in Joint Industry Committees (JICs) in which stakeholders (advertisers, ad companies, media companies) define the metrics (currencies) measuring the size and composition of media audiences. These institutionalized media audiences — “the audience as conceptualized, operationalized, and monetized within the marketplace for media audiences (Napoli, 2003, 2011)” (Napoli, 2012, pp. 80–81) — are crucial for programming and publishing decisions, and as a currency for buying and selling media space for campaigns. Industry data is usually collected in syndicated research by companies such as Nielsen, Arbitron, comSCORE, and Kantar Media. Table 2.2 describes the early days of this type of research in the USA.

The main interest in institutionalized audience measurement has always been exposure (e.g., Ang, 1991; den Boon & Neijens, 2000; Napoli, 2012; Ross & Nightingale, 2003). However, it is clear that the value of media for a campaign depends not only on the number of consumers who come into contact with the medium, but also on the “quality” of these contacts. Therefore, quality indicators such as prestige, communication power, appreciation, involvement and engagement with the medium (content), visibility (of outdoor posters) and so on, and so forth, have always been on the research agenda of the industry, as additions to the mainstream "ratings approach" to audience measurement, even more so in today's interactive media landscape that facilitates these kind of measures (e.g., BRO, 2024; den Boon & Neijens, 2000; Harris & Chasin, 2006; Napoli, 2011, 2012; Smit & Neijens, 2011; Webster et al., 2014).

Table 2.2 The early history of USA rating research

The early history of USA rating research was described by Mayer in 1972: "It started in March 1930, when a group of national advertisers met together with opinion researcher Archibald Crossley to form the Cooperative Analysis of Broadcasting. The technique employed was the 'telephone recall' — interviewers would call people and ask them to run down a list of what they had heard yesterday. The first improvement on this system was Claude E. Hooper's Hooperratings. His technique was the 'telephone-coincidental', in which an interviewer calls a home and asks the person answering the telephone whether the radio (or television set) is playing, and if so what's on" (Mayer, 1972, p. 5). Later, national TV ratings were measured with a combination of a self-administered questionnaire and a household meter. Viewers in the Nielsen Television Index sample kept a diary of their viewing. Those viewing levels were calibrated to ones produced by the Nielsen audiometer, a device attached to televisions, which electronically audited when the machine was turned on and which channels were tuned (Stoddard, 1987). In local and cable markets, only diaries were used, as household meters would have been too costly. This system was replaced in the mid '80s by 'People meters': each television set was fitted with a meter that included a set top comprising a station and channel selection monitor and a display screen, plus a portable remote control handset. Panel members indicate their viewing by pressing pre-designated buttons on the handset (Kent, 1994).

Source: Smit and Neijens (2011, p. 127). Webster et al. (2014) offer an extensive overview of the developments (e.g., institutions, metrics, research designs) of institutionalized audience research.

Two issues have long puzzled the sector: the measurement of small media outlets and cross media measurement. The first issue: the sheer number of media vehicles and the highly fragmented media landscape require huge sample sizes for estimating the size of media audiences. The long tail (Anderson, 2006; Elberse, 2013; McDonald, 2008) indicates the situation in which a very small number of media vehicles or programs have a substantial share of the media market, and many media have small shares that in the aggregate can be larger than the share of the few large players. "It is simply impossible for measurement firms to recruit and maintain representative audience panels that are large enough to capture the true distribution of audience attention across the wealth of available content options and across all of the platforms via which that content can be consumed" (Napoli, 2012,

p. 82). The industry is constantly trying to find solutions, for example using digital trace data (e.g., Smit & Neijens, 2011).

The second issue: the media industry traditionally collects data and produces metrics for individual media types, i.e., for television, print media, outdoor, etcetera. However, the market is no longer satisfied with this "silo" approach and demands cross-media data that contains statistics that also show the overlap of audiences of the different media types. This requires "single source" data that captures all media use of an individual in one survey — (almost) impossible — or data fusion, i.e., linking data sources by matching respondents on common variables, such as demographics, attitudes, and behavior (Jephcott & Bock, 1998).

The industry has taken a large number of initiatives to address the above-mentioned issues, which have become quite complex (Smit & Neijens, 2011; Taneja & Mamoria, 2012). We illustrate this with an approach recently developed in the Netherlands which has not yet been fully implemented (see Table 2.3).

Table 2.3 Advanced cross media audience measurement

The Dutch "Nationaal Media Onderzoek" (National Media Research) aims to document what people watch, listen to, read and click on. For the data collection several instruments are used including app, people meter, router meter, questionnaires, and census data in the following way: (1) in panel number 1, a media app (MediaCell+) is installed on the smartphone of the panel members that records radio listening, tv viewing, and online traffic on the smartphone (using image and audio matching, and an inaudible code embedded in the audio and video audio stream programming to recognize channels and programs), (2) in panel number 2, a household people meter is installed that maps TV viewing behaviour in more detail, and a router meter is installed that registers the IP traffic through the Wi-Fi network in the household, (3) census data records stream starts, downloads and website visits, (4) a readership survey registers readership of print media that combined with the online measurement in the online panel gives data on reading of media brands. The survey also contains questions about respondents' characteristics, information on product and brand cognitions, attitudes and behaviour, and other media use (global). The fifth data source consists of a panel on radio audience measurement (e-diary for people 13–17; it is illegal to track them with an app). An "establishment survey" ties the various data sources together. Outdoor advertising is not (yet) included in the National Media Research.

Source: Nationaal Media Onderzoek (2021).

2.5 Metrics

Based on various sources, we have made an overview of the concepts (also called metrics), used to characterize exposure to media (see Table 2.4).

A distinction can be made between metrics that characterize media usage of individuals (e.g., exposure, attention, frequency of reading), and metrics that characterize media (readership, click-through rate, rating). A distinction within the latter group is between gross measures and cumulative measures (Webster et al., 2014). The first "include estimates of audience size and composition made at a single point in time" (p. 88); the latter show "how audience members are behaving across time" (p. 113). Examples of the first type are market share and impressions. Cumulative measures include cum ratings, reach, unique visitors, frequency, and audience duplication.

An important category of metrics refers to campaigns. Media audiences are usually confronted with a campaign message several times (in the same medium or in different media or on different platforms) which is reflected in metrics such number of contacts, gross reach / gross rating points (GRPs), net reach, duplication, frequency of contacts, and average contact frequency.

Note that there are multiple variants of reach and rating metrics: these can refer to the entire population or a specific target group, to specific times/periods; expressed as an average or share of market ("share of voice"), and may include delayed viewing (time shifting, for instance up to three days).

Data for these metrics can be based on site-centric data (also platform centric, server centric or census data) or user-centric data (Webster et al., 2014; Ohme et al., 2023). An example of site centric data is data from servers that record website visits. Site-centric data is attractive in the fragmented media landscape where sample sizes of surveys restrict the measurement of media outlets with small audiences (see above). Site-centric data, however, is often "data in vacuum" and lack background information about the user. User-centric data is acquired in a panel of users and can be utilized to characterize users as well as media outlets; site-centric data can only be applied to the latter.

Discussions and research on the best way of operationalizing and measuring media metrics are widespread (see the proceedings of the biannual Worldwide Readership Research symposia (e.g., Brown, 1990 and www.readershipsymposium.com). For example, to estimate media readership, respondents may be asked about reading a newspaper or magazine in the last publication interval, a method known as Recent Reading, or about reading specific issues (whenever), a method known as Through-The-Book (Hobson, 1956) or Specific Issue Readership (Faasse & van Meerem, 2003). We return to this in the Chapter 4.

2.6 Concluding remarks

Several concepts are used to describe media contacts. In this book we focus on media exposure (encountering a medium or message), but we also pay attention to psychophysiological reactions such as arousal and valence (Chapters 7 and 8), and mental responses during the encounter, such as attention, involvement, engagement, and experience (Chapter 4) because these aspects are considered important for media use and effects.

This chapter also shows that it is important to consider the medium content (e.g., politics, entertainment, violence, valence), other characteristics of the medium (e.g., device, application, message), and the type of interaction (e.g., active, passive, relational, self-expression) when comparing and evaluating theories and studies of media use and effects.

Table 2.4 Media exposure metrics

Ad exposure	The number of target audience individuals exposed to the advertising (Rossiter & Percy 1998, p. 447).
Associated score	The percentage of respondents who have read some portion of the ad and remembered the name of the advertiser (i.e., brand recognition) (Starch, 1966; Sar & Rodriguez, 2017).
Attention	E.g., How much attention do you pay to news on TV/newspaper articles about national politics?
Average frequency	A campaign generates a frequency distribution of exposures. Average frequency is the mean value of this distribution. See also below (entry: Frequency distribution of contacts).
Average Issue Readership (AIR)	The number of different people reached by an (average) issue of a particular publication.
Campaign reach	The number of contacts with the various messages in the campaign.
Circulation	"Total number of unduplicated audience members exposed to a media vehicle (e.g., newspaper, station) over some specified period of time" (Webster et al., 2014, p. 267).
Click-through rate	The percentage of people who click on a stimulus.
Contact	Example: the number of contacts with a campaign is two if one person is confronted with the campaign twice, or if two persons are confronted with the campaign once.
Cumulative persons/rating	"The total number/percentage of different people or households who have tuned in to a station at least once over some period of time" (Webster et al., 2014, p. 147).
Current reach	The number of different people reached within the publication interval of a title (a day for dailies, a week for weeklies, etc.).
Duplication	The percentage of people who are exposed to media content *x* who are also exposed to media content *y*.

Effective reach	Indicates the number of target audience individuals reached at the effective frequency level, that means the level at which the message has its intended effect.
Exposure	"The extent to which an audience member has encountered specific messages or classes of messages/media content" (Slater, 2004, p. 168).
Frequency distribution of contacts	A campaign generates a frequency distribution of contacts. In a campaign that aims to reach people three times, some individuals will have one, two, three or seven contacts.
Frequency of reading	Respondents are asked how many of the last, say, four or five issues of a publication they have read or looked at.
Gross reach	The total number of contacts with a campaign.
GRP	Gross Rating Point. Example: a value of 100 GRPs means that the number of contacts are equal to the number of targeted individuals. It may mean that 50% of the target individuals are reached twice. Gross Rating Points: net reach * average frequency.
Impressions	The number of times that a message is displayed to a user.
Net reach	The number of different people who are exposed to a campaign at least once.
Noted scores	The percentage of respondents who remembered seeing a specific ad (i.e., ad recognition) (Starch, 1966; Sar & Rodriguez, 2017).
Opportunity to see (OTS)	Open eyes/ears in front of medium space (Slater, 2004).
Overlap	Gross reach minus net reach = overlap (repeated exposures to a campaign).
Page views	The number of times a page has been loaded.
Rating	The percentage of a population exposed to a (broadcast) medium.
Reach	The number of different people who are exposed to a medium or a message. Similar to cumulative persons/rating.
Read most	The percentage of respondents who have read more than 50% of the ad (Starch, 1966; Sar & Rodriguez, 2017).
Recent Reading	Determines whether a respondent has read (or claims to have read) any copy of a publication in the most recent publishing interval (the past month for monthlies, the past week for weeklies, etc.).
Selectivity	The number of target group people reached compared to the total number of people reached.
Screen time (also, screen use, (digital) media time, (digital) technology use, digital engagement)	"Time spent using a device such as a computer, television, or games console" (The *Oxford English Dictionary*; Kaye et al., 2020). "Time spent passively watching screen-based entertainment (TV, computer, mobile devices)" (World Health Organization; Kaye et al., 2020).
Share	The number of people who are, for instance, tuned to channel *x* or read magazine *y* divided by the total number of people tuned to television/reading magazines.

Subscriptions	The number of subscriptions (paid, for free, controlled circulation, etc.).
TRP	Target Rating Points: gross rating points for a targeted audience.
Unduplicated audience	Similar to cumulative persons/rating and reach (see entry above).
Unique visitors	The number of different people who have visited a website over some period of time.
Waste	The number of people reached who do not belong to the target group.

Literal quotes are indicated in the table with "…".
General sources that form the basis of the other descriptions are: Kaye et al. (2020); Media Rating Council [MRC] and Interactive Advertising Bureau [IAB] (2013); Moorman et al. (2011); Rossiter and Percy (1998); Webster et al. (2014).

References

Advertising Research Foundation (1961). *Toward better media comparisons*. Advertising Research Foundation.

Advertising Research Foundation (2006a). *Engagement: Definitions and anatomy*. Advertising Research Foundation.

Allen, R. L., & Waks, L. (1986, May 22–26). *The dimensionality of commonly employed mass media exposure measures* [Paper presentation]. International Communication Association Annual Conference, Chicago, IL, USA.

Ang, I. (1991). *Desperately seeking the audience*. Routledge.

Anderson, C. (2006). *The long tail: Why the future of business is selling less of more*. Hachette Books.

Araujo, T. B., Copulsky, J. R., Hayes, J. L., Kim, S. J., & Srivastava, J. (2020). From purchasing exposure to fostering engagement: Brand-consumer experiences in the emerging computational advertising landscape. *Journal of Advertising*, *49*(4), 428–445. https://doi.org/10.1080/00913367.2020.1795756

Bellman, S., Nenycz-Thiel, M., Kennedy, R., Hartnett, N., & Varan, D. (2019). Best measures of attention to creative tactics in TV advertising: When do attention-getting devices capture or reduce attention?. *Journal of Advertising Research*, *59*(3), 295–311. https://doi.org/10.2501/JAR-2019-002

BRO (2024). *BRO Next video*. https://buitenreclame-onderzoek.nl/video-uk

Bronner, F., & Neijens, P. (2006). Audience experiences of media context and embedded advertising: A comparison of eight media. *International Journal of Market Research*, *48*(1), 81–100. https://doi.org/10.1177/147078530604800106

Brown, M. (1990). *Dear reader: Some readership measurement questions ... and some answers*. Research Services Ltd & the British Market Research Bureau Ltd.

Calder, B. J., Isaac, M. S., & Malthouse, E. C. (2016). How to capture consumer experiences: A context-specific approach to measuring engagement predicting consumer behavior across qualitatively different experiences. *Journal of Advertising Research*, *56*(1), 39–52. https://doi.org/10.2501/JAR-2015-028

Chaffee, S. H., & Schleuder, J. (1986). Measurement and effects of attention to media news. *Human Communication Research*, *13*(1), 76–107. https://doi.org/10.1111/j.1468-2958.1986.tb00096.x

Csikszentmihalyi, M. (1988). The flow experience and its significance for human psychology. In M. Csikszentmihalyi & I. S. Csikszentmihalyi (Eds.), *Optimal experience: Psychological studies of flow in consciousness* (pp. 15–35). Cambridge University Press.

den Boon, A.K., & Neijens, P. C. (Eds.). (2000). *Media en reclame. Mediaplanning en bereiksonderzoek* (3rd ed.). Wolters-Noordhoff.

Elberse, A. (2013). *Blockbusters: Hit-making, risk-taking, and the big business of entertainment*. Holt.

Faasse, J., & van Meerem, L. (2003, October 26–29). *What the world needs now* [Paper presentation]. Worldwide Readership Research Symposium, Cambridge, UK.

Fikkers, K. M., Piotrowski, J. T., & Valkenburg, P. M. (2015). Assessing the reliability and validity of television and game violence exposure measures. *Communication Research*, *44*(1), 117–143. https://doi.org/10.1177/0093650215573863

Finn, A. (1988). Print ad recognition readership scores: An information processing perspective. *Journal of Marketing Research*, *25*(2), 168–177. https://doi.org/10.2307/3172648

Geiger, S., & Newhagen, J. (1993). Revealing the black box: Information processing and media effects. *Journal of Communication*, *43*(4), 42–50. https://doi.org/10.1111/j.1460-2466.1993.tb01303.x

Grady, S. M., Tamborini, R., Eden, A., & van der Heide, B. (2022). The social factors and functions of media use. *Journal of Communication*, *72*(5), 523–539. https://doi.org/10.1093/joc/jqac026

Greenwald, A. G., & Leavitt, C. (1984). Audience involvement in advertising: Four levels. *Journal of Consumer Research*, *11*(1), 581–592. https://doi.org/10.1086/208994

Griffioen, N., van Rooij, M. M. J. W., Lichtwarck-Aschoff, A., & Granic, I. (2020b). Toward improved methods in social media research. *Technology, Mind, and Behavior*, *1*(1). https://doi.org/10.1037/tmb0000005

Harris, C. D., & Chasin, J. (2006, June 20–21). *The impact of technological innovation on media exposure tracking: In search of "the new traditional"* [Paper presentation]. Advertising Research Foundation Audience Measurement Symposium, NY, USA.

Harvey, B. (1997). The expanded ARF model: Bridge to the accountable advertising future. *Journal of Advertising Research*, *37*(2), 11–20.

Higgins, E. T., & Scholer, A. A. (2009). Engaging the consumer: The science and art of the value creation process. *Journal of Consumer Psychology*, *19*(2), 100–114. https://doi.org/10.1016/j.jcps.2009.02.002

Hobson, J. W. (1956). *The selection of advertising media* (2nd ed.). Business Publications on behalf of the Institute of Practitioners in Advertising.

Jephcott, J., & Bock, T. (1998). The application and validation of data fusion. *Journal of the Market Research Society*, *40*(3), 1–18. https://doi.org/10.1177/147078539804000301

Johnson, B. T., & Eagly, A. H. (1989). Effects of involvement on persuasion: A meta-analysis. *Psychological Bulletin*, *106*(2), 290–314. https://doi.org/10.1037/0033-2909.106.2.290

Katz, E., Blumler, J. G., & Gurevitch, M. (1973). Uses and gratifications research. *Public Opinion Quarterly*, *37*(4), 509–523. https://doi.org/10.1086/268109

Kaye, L. K., Orben, A., Ellis, D. A., Hunter, S. C., & Houghton, S. (2020). The conceptual and methodological mayhem of "Screen Time." *International Journal of Environmental Research and Public Health*, *17*(10), Article 3661. https://doi.org/10.3390/ijerph17103661

Kent, R. (1994). Measuring media audiences. An overview. In R. Kent (Ed.), *Measuring media audiences*. Routledge.

Koppe, I. (1998). Literature study media experience. In I. Koppe (Ed.), *Theme dossier media experience* (pp. III-1–III-109). Platform, 95.

Lee, J. K. (2009). *Incidental exposure to news: Limiting fragmentation in the new media environment* [Doctoral dissertation, University of Texas at Austin]. UT Electronic Theses and Dissertations. http://hdl.handle.net/2152/6686

Malthouse, E. C., Calder, B. J., & Eadie, W. P. (2003). *Conceptualizing and measuring magazine reader experiences*. Media Management Center, Northwestern University.

Malthouse, E. C., Calder, B. J., & Tamhane, A. (2007). The effects of media context experiences on advertising effectiveness. *Journal of Advertising*, *36*(3), 7–18. https://doi.org/10.2753/JOA0091-3367360301

Mayer, M. (1972). A writer looks at TV ratings. *Journal of Advertising Research*, *12*(4), 3–10.

McDonald, S. (2008). The long tail and its implications for media audience measurement. *Journal of Advertising Research*, *48*(3), 313–319. http://doi.org/10.2501/S0021849908080379

McQuail, D. (1994). *Mass communication theory: An introduction* (3rd ed.). Sage.

Media Rating Council [MRC], & Interactive Advertising Bureau [IAB]. (2013, August). *Audience reach measurement guidelines version 2.0—Updated August 2013*. IAB. https://www.iab.com/wp-content/uploads/2015/06/AudienceReachMeasurementGuidelines.pdf

Meier, A., & Reinecke, L. (2021). Computer-mediated communication, social media, and mental health: A conceptual and empirical meta-review. *Communication Research*, *48*(8), 1182–1209. https://doi.org/10.1177/0093650220958224

Moorman, M., Neijens, P. C., & Smit, E. G. (2005). The effects of program responses on the processing of commercials placed at various positions in the program and the block. *Journal of Advertising Research*, *45*(1), 49–59. https://doi.org/10.1017/S0021849905050105

Moorman, M., Neijens, P. C., & Smit, E. G. (2011). Advertising media selection and planning. In J. Sheth & N. Malhotra (Eds.), *Wiley international encyclopedia of marketing, Vol. 4* (pp. 22–29). Wiley.

Moray, N. (1970). *Attention: Selective processes in vision and hearing*. Academic Press.

Muntinga, D. G., Moorman, M., & Smit, E. G. (2011). Introducing COBRAs. *International Journal of Advertising*, *30*(1), 13–46. https://doi.org/10.2501/IJA-30-1-013-046

Nabi, R. L. (2009). Emotion and media effects. In R. L. Nabi & M. B. Oliver (Eds.), *The Sage handbook of media processes and effects* (pp. 205–221). Sage.

Nabi, R. L., & Oliver, M. B. (Eds.). (2009). *The Sage handbook of media processes and effects*. Sage.

Napoli, P. M. (2003). *Audience economics: Media institutions and the audience marketplace*. Columbia University Press.

Napoli, P. M. (2011). *Audience evolution: New technologies and the transformation of media audiences*. Columbia University Press.

Napoli, P. M. (2012). Audience evolution and the future of audience research. *International Journal on Media Management*, *14*(2), 79–97. https://doi.org/10.1080/14241277.2012.675753

Nationaal Media Onderzoek (2021). *Nationaal Media Onderzoek* (*EN-version*) [Video]. YouTube. https://www.youtube.com/watch?v=PMInIkiFAms

Ohme, J., Araujo, T., Boeschoten, L., Freelon, D., Ram, N., Reeves, B. B., & Robinson, T. N. (2023). Digital trace data collection for social media effects research: APIs, data donation, and (screen) tracking. *Communication Methods and Measures*. Advance online publication. https://doi.org/10.1080/19312458.2023.2181319

Oliver, M. B., Raney, A. R., & Bryant, J. (Eds.). (2020). *Media effects: Advances in theory and research*. Routledge.

Parry, D., Davidson, B. I., Sewall, C., Fisher, J. T., Mieczkowski, H., & Quintana, D. S. (2021). A systematic review and meta-analysis of discrepancies between logged and self-reported digital media use. *Nature Human Behavior*, *5*(11), 1535–1547. https://doi.org/10.1038/s41562-021-01117-5

Petty, R. E., Cacioppo, J. T., & Schumann, D. (1983). Central and peripheral routes to advertising effectiveness: The moderating role of involvement. *Journal of Consumer Research*, *10*(2), 135–146. https://doi.org/10.1086/208954

Pieters, R. F. G. M., & Wedel, M. (2004). Attention capture and transfer in advertising: Brand, pictorial, and text-size effects. *Journal of Marketing*, *68*(2), 36–50. https://doi.org/10.1509/jmkg.68.2.36.27794

Pieters, R. F. G. M., & Wedel, M. (2007). Goal control of attention to advertising: The Yarbus implication. *Journal of Consumer Research*, *34*(2), 224–233. https://doi.org/10.1086/519150

Posner, M. I., & Boies, S. J. (1971). Components of attention. *Psychological Review*, *78*(5), 391–408. https://doi.org/10.1037/h0031333

Potter, W. J. (2008). The importance of considering exposure states when designing survey research studies. *Communication Methods and Measures*, *2*(1–2), 152–166. https://doi.org/10.1080/19312450802062299

Potter, W. J. (2011). Conceptualizing mass media effect. *Journal of Communication*, *61*(5), 896–915. https://doi.org/10.1111/j.1460-2466.2011.01586.x

Potter, W. J. (2012). *Media effects*. Sage.

Price, V., & Zaller, J. (1993). Who gets the news? Alternative measures of news reception and their implications for research. *Public Opinion Quarterly*, *57*(2), 133–64. https://doi.org/10.1086/269363

Romantan, A., Hornik, R., Price, V., Cappella, J., & Viswanath, K. (2008). A comparative analysis of the performance of alternative measures of exposure. *Communication Methods and Measures*, *2*(1–2), 80–99. https://doi.org/10.1080/19312450802062539

Rosbergen, E., Pieters, R. F. G. M., & Wedel, M. (1997). Visual attention to advertising: A segment-level analysis. *Journal of Consumer Research*, *24*(3), 305–314. https://doi.org/10.1086/209512

Ross, K., & Nightingale, V. (2003). *Media and audiences: New perspectives*. Open University Press.

Rossiter, J. R., & Percy, L. (1998). *Advertising communications and promotion management*. McGraw-Hill.

Russell, J. A. (1980). A circumplex model of affect. *Journal of Personality and Social Psychology*, *39*(6), 1161–1178. https://doi.org/10.1037/h0077714

Sar, S., & Rodriguez, L. (2017). Copy test and Starch test. In J. Matthes, C. S. Davis, & R. F. Potter (Eds.), *The international encyclopedia of communication research methods*. Wiley & Sons. https://doi.org/10.1002/9781118901731.iecrm0047

Shapiro, S., MacInnis, D. J., & Heckler, S. E. (1997). The effects of incidental ad exposure on the formation of consideration sets. *Journal of Consumer Research*, *24*(1), 94–104. https://doi.org/10.1086/209496

Slater, M. D. (2004). Operationalizing and analyzing exposure: The foundation of media effects research. *Journalism & Mass Communication Quarterly*, *81*(1), 168–183. https://doi.org/10.1177/107769900408100112

Smit, E. G., & Neijens, P. C. (2011). The march to reliable metrics: A half-century of coming closer to the truth. *Journal of Advertising Research*, *51*(1), 124–135. https://doi.org/10.2501/JAR-51-1-124-135

Smit, E. G., Neijens, P. C., & Heath, R. (2013). The differential effects of position, ad and reader characteristics on readers' processing of newspaper ads. *International Journal of Advertising*, *32*(1), 65–84. https//doi.org/10.2501/IJA-32-1-065-084

Starch, D. (1966). *Measuring advertising readership and results*. McGraw-Hill Book Company.

Stoddard Jr., L. R. (1987). The history of people meters: How we got to where we are (and why). *Journal of Advertising Research*, *27*(5), 10–12.

Taneja, H., & Mamoria, U. (2012). Measuring media use across platforms: Evolving audience information systems. *The International Journal on Media Management*, *14*(2), 121–140. https://doi.org/10.1080/14241277.2011.648468

Valkenburg, P. M., & Peter, J. (2013a). Five challenges for the future of media-effects research. *International Journal of Communication*, *7*, 197–215. http://ijoc.org/index.php/ijoc/article/download/1962/849

van der Wal, A., Valkenburg, P. M., & van Driel, I. I. (2022). *In their own words: How adolescents differ in their social media use and how it affects them*. PsyArXiv. https://doi.org/10.31234/osf.io/mvrpn

Webster, J. G., Phalen, P. F., & Lichty, L. W. (2014). *Ratings analysis: Audience measurement and analytics* (4th ed.). Routledge.

Webster, J. G., & Wakshlag, J. (1985). Measuring exposure to television. In D. Zillman & J. Bryant (Eds.), *Selective exposure to communication* (pp. 35–62). Routledge.

Weinstein, E. (2018). The social media see-saw: Positive and negative influences on adolescents' affective well-being. *New Media & Society*, *20*(10), 3597–3623. https://doi.org/10.1177/1461444818755634

Yang, C., Holden, S. M., & Ariati, J. (2021). Social media and psychological well-being among youth: The multidimensional model of social media use. *Clinical Child and Family Psychology Review*, *24*(3), 631–650. https://doi.org/10.1007/s10567-021-00359-z

Zaichkowsky, J. L. (1985). Measuring the involvement construct. *Journal of Consumer Research*, *12*(3), 341–352. http://www.jstor.org/stable/254378

About the authors

Peter Neijens is an Honorary Fellow and Emeritus Professor at ASCoR. He held the Chair in Media and Persuasion at the University of Amsterdam until 2019. His research focuses on persuasive communication, in particular media use and effects in advertising and public opinion.

Theo Araujo is a Full Professor of Media, Organisations and Society, and Scientific Director of ASCoR. His research investigates the dynamic interplay between media and organizations, and what it means for society, with a special focus on trust and technology. His research has also a methodological focus on computational communication science and the implementation of large-scale data collection and analysis for communication research.

Judith Möller holds the chair of "Empirical Communication Research, especially Media Use and Social Media Effects" at the University of Hamburg and the Leibniz Institute for Media Research | Hans Bredow Institute (HBI). Before 2023 she was an Associate Professor of Political Communication at the Amsterdam School of Communication Research.

Claes H. de Vreese is a Distinguished University Professor of AI and Society with a special emphasis on media and democracy. He co-directs the Gravitation Program AlgoSoc and the AI, Media and Democracy Lab. He holds the chair in Political Communication at the Amsterdam School of Communication Research.

3. Quality Criteria for Media Exposure Measures

Abstract
Quality criteria for media exposure measures can be derived from the Total Survey Error (TSE) paradigm that specifies "all errors that may arise in the design, collection, processing, and analysis of survey data" (Biemer, 2010, p. 817; Groves, 1989; Groves et al., 2009; Weisberg, 2015).[1] Roughly speaking, we can distinguish two kinds of errors: measurement errors (differences between the observed and true values of variables) and respondent selection errors (differences between the target and participating population). These are discussed in the context of MCE measurement.

Keywords: reliability, validity, sampling errors, non-response, platform affordances errors, trace selection errors

3.1 Measurement errors: reliability and validity criteria

Reliability and validity are of crucial importance for the measurement of MCE: "A poor reliability of exposure measures can drastically attenuate the relationship with outcome variables, and a low validity makes it difficult to interpret any relationship (Fern & Monroe, 1996; McGuire, 1986)" (Valkenburg & Peter, 2013a, p. 200). Table 3.1 lists a number of reliability and validity indicators used in studies on the quality of exposure measures.

Compared to the number of validity studies, the number of reliability studies is limited (Lee et al., 2008). Test-retest reliability is probably the most applied reliability assessment strategy for media exposure

1 A broader concept is total survey quality (e.g., Juran & Gryna, 1980) that also includes "fitness for use" dimensions such as timeliness, accessibility, relevance, and usability of the data (Biemer, 2010, p. 818). In the context of media exposure data, "scalability" is such a criterion: the extent to which a measure can be applied to a large number of respondents.

Neijens, P., Araujo, T., Möller, J., & de Vreese, C. H. (2024). Quality criteria for media exposure measures. In P. Neijens, T. Araujo, J. Möller, & C. H. de Vreese, *Measuring exposure and attention to media and communication. Solutions to wicked problems* (pp. 33-40). Amsterdam University Press. https://doi.org/10.5117/9789463723176_ch03

Table 3.1 Criteria for reliability and validity

Criterion	*Definition/operationalization*
Reliability	"The extent to which an experiment, test, or any measuring procedure yields the same results on repeated trials" (Carmines & Zeller, 1979, p. 11). Also: "Reliability refers to the notion that re-elicitation of the same measure from a respondent should produce an identical response as long as the underlying true exposure is unchanged" (Lee et al., 2008, p. 6).
Internal consistency	Cronbach's alpha (applicable to a case with different questions, same construct; cannot be applied to single items).
Test-retest reliability	Consistency of the scores when applying the same measure several times.
Validity	"The extent to which a test accurately reflects or assesses the specific construct it purports to measure" (Coolidge & Segal, 2010, p. 1).
Content validity	"The extent to which an empirical measurement reflects a specific domain of content" (Carmines & Zeller, 1979, p. 20).
Face validity	"Concerns judgements about an instrument after it is constructed" (Nunnally, 1978, p. 111). "The extent to which it looks like it measures what it is intended to measure" (Carmines & Zeller, 1979, p. 53).
Nomological validity	The extent to which a measure is positively associated with other variables that theory predicts that the measure ought to be associated with (compare Niederdeppe, 2014, p. 148).
Predictive validity	The extent to which a measure is positively associated with a presumed outcome variable at a later time, e.g., a *change* in a dependent variable, such as *change* in knowledge.
Convergent/concurrent validity	The extent to which a measure is associated with alternative measures.
Criterion validity	The extent to which a measure is associated with a well-established measure (gold standard).
Discriminant validity	The extent to which a measure is unrelated to unrelated concepts.
Discriminant validity: Accuracy	"The extent to which campaign exposure measures successfully distinguish between 'true' campaign messages (those that a respondent could have been exposed to based on availability) and 'false' ones (those that were not available such that respondents could not have been exposed to them" (Niederdeppe, 2014, p. 147).
Discriminant validity: sensitivity	"The extent to which a campaign measure is distinct from broader measures of media use or exposure to other sources of topical information" (Niederdeppe, 2014, p. 147).

measures. It compares the consistency of the scores when applying the same test two or more times, for example in panel studies. Statistical issues associated with this approach are discussed in Fikkers et al. (2015), Heise (1969), Lee et al. (2008), Prior (2013), Wiley and Wiley (1970), and Zaller (2002).

In addition to the "soft" face validity criterium, three main validity criteria are predictive or nomological validity,[2] convergent or criterion validity (both criteria compare measures, the difference is that whether the comparison variable can be considered a gold standard (criterion) or not), and discriminant validity.

A predictive validity approach is applied in studies that compare media exposure measures with changes in a variable which is supposed to be affected by the exposure, e.g., political knowledge. It is important that these analyses control for alternative time-variant causes of the selected variable (i.e., political knowledge). It is also important that these studies make it plausible that the media to which these people are exposed actually contain information about the knowledge measured in these questions (at the time when the respondents consumed these media) and that they were able to obtain and retain this information (Prior, 2013).

Panel studies are instrumental for test-retest reliability and predictive validity studies. Another interesting study design are experiments in which respondents in some conditions are exposed to a message (criterion), and others not, and asked to self-report their exposure to the message (e.g., Jerit et al., 2016; Vavreck, 2007). Some studies consider digital trace data or diary data the gold standard for self-reports, but given the issues associated with these "gold standards" this is certainly not widely accepted (see Chapter 5).

Please note that the association between (media exposure) variables does not tell us to what extent the scores on the variables are similar in terms of degree: it is possible that people overestimate their exposure to media measured with a particular instrument, while at the same time this measure correlates perfectly with a gold standard.

Associations between different exposure measures (in other words: estimations of convergent or criterion validity), or between media exposure and a dependent variable (estimations of predictive validity) can be studied at a "between-person" or a "within-person" level (Valkenburg et al., 2021; Verbeij et al., 2022). Between-person analysis examines these associations at the group level: e.g., it is investigated to what extent respondents who score higher on a media exposure measure than other respondents, also score higher on (for instance) knowledge increase than the other respondents. Within-person analysis requires panel data and examines these associations

2 Niederdeppe (2014) prefers the label nomological rather than predictive validity "because the latter term often implies evidence of causal ordering and the use of longitudinal data" (p. 148).

at the individual level: it is investigated to what extent people who score higher/lower on a media exposure measure than usual also score higher/lower on (for instance) knowledge increase than they usually score (resulting in a person specific, "$n = 1$" beta) (Valkenburg et al., 2024; Verbeij et al., 2022).

Coding of content

Many of the exposure measures obtained via self-reports or digital trace data (see Chapters 4 and 5) may require manual coding (i.e., labeling) of the content into broader categories especially when relying on very granular levels of exposure measurement. For example, a list of websites or domains visited by a participant may need to be recoded into a broader category of content types (e.g., news, sports, entertainment, etc.). In some cases, especially in cases with extremely large datasets, a machine learning classifier may be trained on the manually coded data (a sample of the larger dataset), and subsequently used to automatically categorize the complete dataset. Additional reliability measures are then needed to ensure the quality of such coding, as outlined in Table 3.2.

Table 3.2 Criteria for reliability for (manual) coding

Criterion	*Definition/operationalization*
Intercoder reliability	"Reliability is measured by the agreement among all pairs of independent coders who are instructed to identify, interpret and record the same set of units of analysis. This agreement measure indicates the extent to which the process of generating data is reproducible elsewhere, by additional coders or for other applications" (Krippendorff, 2021, p. 166).
Intercoder reliability measures	Traditionally used for content analysis and applicable to: same test/task, different coders; include measures such as Krippendorff's alpha.
Information retrieval measures	Measures such as Precision, Recall or F1-Scores to assess the performance of machine learning predictions, relevant for when manual coding of a sample is used as an input to train a classifier that will be used to make predictions (i.e., categorize) the full dataset.

3.2 Respondent selection errors: frame, sampling and non-response issues

Respondent selection errors include frame (coverage) errors, sampling, and non-response issues. Coverage is at risk if the sampling frame does not coincide with the population of interest. Sampling is a problem if the sampling scheme and/or sample size does not allow for an accurate picture of the population of interest. Non-response problems occur when respondents from the sample are not available, or refuse to participate, the probability of which increases with more intensive MCE measurement methods. Response quality in surveys is not only threatened by people who refuse to participate in a study (unit non-response), but also by research participants who do not answer (all) questions in the study (item non-response or non-compliance).

Inspired by the TSE framework, Sen et al. (2021) introduced a conceptual framework to diagnose, understand, and document errors that are specific to the use of digital traces (the "Total Error Framework for Digital Traces of Human Behavior on Online Platforms") including error types such as platform coverage error (the extent to which platform users deviate from the target population), platform affordances error (the extent to which the affordances of a platform influence the digital traces), trace selection error (the extent to which queries fail to capture all relevant posts or include irrelevant posts), and user selection error (the gap between the type of users selected and the type of users comprising the platform's userbases).

3.3 Concluding remarks

Quality criteria for media exposure measurements can be derived from the TSE paradigm, which includes sampling (or representation) errors (e.g., non-response, selection) and measurement errors (e.g., errors arising from survey questions or trace selection). Response, reliability, and validity are therefore important criteria for the quality of (exposure) measurements, and will be used as a guideline to understand the advantages and limitations of the different MCE measurement strategies in the following chapters.

References

Biemer, P. P. (2010). Total survey error: Design, implementation, and evaluation. *Public Opinion Quarterly*, *74*(5), 817–848. https://doi.org/10.1093/poq/nfq058

Carmines, E. G., & Zeller, R. A. (1979). *Reliability and validity assessment*. Sage.

Coolidge, F. L., & Segal, D. L. (2010). Validity. In I. B. Weiner & W. E. Craighead (Eds.), *The Corsini encyclopedia of psychology: Vol. 4*. John Wiley & Sons. https://doi.org/10.1002/9780470479216.corpsy1019

Fern, E. F., & Monroe, K. B. (1996). Effect-size estimates: Issues and problems in interpretation. *Journal of Consumer Research*, *23*(2), 89–105. https://doi.org/10.1086/209469

Fikkers, K. M., Piotrowski, J. T., & Valkenburg, P. M. (2015). Assessing the reliability and validity of television and game violence exposure measures. *Communication Research*, *44*(1), 117–143. https://doi.org/10.1177/0093650215573863

Groves, R. (1989). *Survey errors and survey costs*. John Wiley & Sons.

Groves, R. M., Fowler Jr., F. J., Couper, M. P., Lepkowski, J. M., Singer, E., & Tourangeau, R. (2009). *Survey methodology* (2nd ed.). Wiley.

Heise, D. R. (1969). Separating reliability and stability in test-retest-correlations. *American Sociological Review*, *34*(1), 93–101. https://doi.org/10.2307/2092790

Jerit, J., Barabas, J., Pollock, W., Banducci, S., Stevens, D., & Schoonvelde, M. (2016). Manipulated vs. measured: Using an experimental benchmark to investigate the performance of self-reported media exposure. *Communication Methods and Measures*, *10*(2–3), 99–114. https://doi.org/10.1080/19312458.2016.1150444

Juran, J., & Gryna, F. (1980). *Quality planning and analysis* (2nd ed.). McGraw-Hill.

Krippendorff, K. (2021). A quadrilogy for (big) data reliabilities. *Communication Methods and Measures*, *15*(3), 165–189. https://doi.org/10.1080/19312458.2020.1861592

Lee, C., Hornik, R., & Hennessy, M. (2008). The reliability and stability of general media exposure measures. *Communication Methods and Measures*, *2*(1–2), 6–22. http://dx.doi.org/10.1080/19312450802063024

McGuire, W. J. (1986). The myth of massive media impact: Savagings and salvagings. In G. Comstock (Ed.), *Public communication and behaviour, Vol. 1* (pp. 173–257). Academic Press.

Niederdeppe, J. (2014). Conceptual, empirical, and practical issues in developing valid measures of public communication campaign exposure. *Communication Methods and Measures*, *8*(2), 138–161. https://doi.org/10.1080/19312458.2014.903391

Nunnally, J. C. (1978). *Psychometric theory*. McGraw-Hill.

Prior, M. (2013). The challenge of measuring media exposure: Reply to Dilliplane, Goldman, and Mutz. *Political Communication*, *30*(4), 620–634. https://doi.org/10.1080/10584609.2013.819539

Sen, I., Flöck, F., Weller, K., Weiß, B., & Wagner, C. (2021). A total error framework for digital traces of human behavior on online platforms. *Public Opinion Quarterly*, *85*(S1), 399–422. https://doi.org/10.1093/poq/nfab018

Valkenburg, P. M., Beyens, I., Bij de Vaate, N., Janssen, L., & van der Wal, A. (2024). Person-specific media effects. In T. Araujo & P. Neijens (Eds.), *Communication research into the digital society: Fundamental insights from the Amsterdam School of Communication Research* (pp. 233–245). Amsterdam University Press. https://doi.org/10.5117/9789048560592_ch14

Valkenburg, P. M., Beyens, I., Pouwels, J. L., van Driel, I. I., & Keijsers, L. (2021). Social media and adolescents' self-esteem: Heading for a person-specific media effects paradigm. *Journal of Communication*, *71*(1), 56–78. https://doi.org/10.1093/joc/jqaa/039

Valkenburg, P. M., & Peter, J. (2013a). Five challenges for the future of media-effects research. *International Journal of Communication*, *7*, 197–215. http://ijoc.org/index.php/ijoc/article/download/1962/849

Vavreck, L. (2007). The exaggerated effects of advertising on turnout: The dangers of self-reports. *Quarterly Journal of Political Science*, *2*(4), 325–343. https://doi.org/10.1561/100.00006005

Verbeij, T., Pouwels, J. L., Beyens, I., & Valkenburg, P. M. (2022). Experience sampling self-reports of social media use have comparable predictive validity to digital trace measures. *Scientific Reports*, *12*(1), Article 7611. https://doi.org/10.1038/s41598-022-11510-3

Weisberg, H. F. (2015). Total survey error. In L. R. Atkeson & R. M. Alvarez (Eds.), *The handbook of polling and survey methods* (pp. 13–27). Oxford University Press.

Wiley, D. E., & Wiley, J. A. (1970). The estimation of measurement error in panel data. *American Sociological Review*, *35*(1), 112–117. https://doi.org/10.2307/2093858

Zaller, J. (2002). The statistical power of election studies to detect media exposure effects in political campaigns. *Electoral Studies*, *21*(2), 297–329. https://doi.org/10.1016/S0261-3794(01)00023-3

About the authors

Peter Neijens is an Honorary Fellow and Emeritus Professor at ASCoR. He held the Chair in Media and Persuasion at the University of Amsterdam until 2019. His research focuses on persuasive communication, in particular media use and effects in advertising and public opinion.

Theo Araujo is a Full Professor of Media, Organisations and Society, and Scientific Director of ASCoR. His research investigates the dynamic interplay between media and organizations, and what it means for society, with a special focus on trust and technology. His research has also a methodological focus on computational communication science and the implementation of large-scale data collection and analysis for communication research.

Judith Möller holds the chair of "Empirical Communication Research, especially Media Use and Social Media Effects" at the University of Hamburg and the Leibniz Institute for Media Research | Hans Bredow Institute (HBI).

Before 2023 she was an Associate Professor of Political Communication at the Amsterdam School of Communication Research.

Claes H. de Vreese is a Distinguished University Professor of AI and Society with a special emphasis on media and democracy. He co-directs the Gravitation Program AlgoSoc and the AI, Media and Democracy Lab. He holds the chair in Political Communication at the Amsterdam School of Communication Research.

4. Self-Report Measures

Abstract

In section 4.1 we discuss the different types of self-reports and their pros and cons. We zoom in on "momentary" or "in situ" self-reports, which have become increasingly popular in contemporary research, in section 4.2. Quality issues of self-reports are discussed in section 4.3. Measures of mental responses during exposure (see Chapter 2) — which are considered important in studies on media use and effects — are discussed in section 4.4. Conclusions about self-reports are drawn in section 4.5. Because self-reports can be formulated in many different ways, and the large number of methodological studies into the quality of self-report variants, this is an extensive chapter.

Keywords*:* recall, recognition, mental responses, momentary/in situ measurement, compliance, validity

4.1 Types of self-report measures of media exposure and their pros and cons

"How many days in a standard week do you watch television?", "And on the days that you watch television, how much time do you spend on this per day?" These questions are typical examples of self-report measures, asking for frequency and duration of media exposure respectively. The advantages of self-report measures (based on media users' own assessments of their media and communication exposure) are obvious: they are easy to include in a questionnaire or diary in which also other questions can be inserted. This makes it possible to correlate media exposure with individual and context characteristics, and with reported cognitions, attitudes, opinions, or behavior. Studies with self-reports of media use are widespread. For instance, in their review of 94 empirical studies on social media use and well-being (period 2010–2018), Griffioen et al. (2020b) found that about

Neijens, P., Araujo, T., Möller, J., & de Vreese, C. H. (2024). Self-report measures. In P. Neijens, T. Araujo, J. Möller, & C. H. de Vreese, *Measuring exposure and attention to media and communication. Solutions to wicked problems* (pp. 41-86). Amsterdam University Press. https://doi.org/10.5117/9789463723176_ch04

82% quantified social media use by asking participants to retrospectively report on their social media use.

Self-reports typically rely on “respondents’ ability to recognize or recall some level of detail of a message or campaign to assess exposure” (Niederdeppe, 2014, p. 142). Different dimensions, such as type of recall, unit of measurement, reference period, response scale, help, and linkage, can be used to distinguish different types of self-report measures (see Table 4.1).

Types of recall

Recall measures were developed by George Gallup in the 1930s, and basically ask media users which media or media messages they recall. Types include free (or unaided) recall, aided (or cued) recall (in which a respondent is provided with a cue), and proven (or confirmed) recall (where the respondent must prove that (s)he has heard or seen a program or an ad) (see examples in Table 4.1).

An alternative to these recall measurers are recognition measures, developed by Daniel Starch in 1932. In his method a picture of a medium message is shown to respondents, who are then asked whether they remembered the message. A well-known example of the recognition method for measuring reach of magazines is the “Through the Book” method (Hobson, 1956), in which respondents are shown a (stripped) issue, or the front page, of a publication, and asked whether they have read or opened the issue before.

An issue with applying recognition, aided and proven recall measures is that photos or other material have to be added to the questionnaire, which raises the question of what should be disclosed (e.g., what level of detail?) and what content respondents should remember (Niederdeppe, 2014). “Aided ad recall measures require the evaluator to develop short ad summaries that are both concise enough for survey measurement but distinct enough to differentiate the message from others” (Niederdeppe, 2014, p. 151). Proven recall measures require content analysis of the open-ended question which can be time consuming and problematic (Slater, 2004).

The (different) meaning of recall and recognition has been the subject of intensive discussions in the literature. This included the extent to which these two measures assess exposure, memory, or interest, and how message type (emotional versus cognitive) and medium type (TV versus print) influence the differences between the scores of the two measures (see Smit & Neijens, 2011 for details). A general finding is that recognition and aided recall usually produce higher estimates of exposure than confirmed recall, which in turn produce higher estimates than unaided recall. These measures correlate moderately to highly in most campaign studies, which implies that convergent validity of these measures is somewhat satisfactory (Niederdeppe, 2014).

Table 4.1 Types of self-reports of media exposure measures (with examples)

Different types of recall Unaided/Free recall – Please list the brands that were advertised in yesterday's newspaper Aided/Cued recall (Researcher provides some detail:) – Have you seen an ad for *x* that showed … – Which of the following programs do you regularly watch on television? Please check any that you watch at least once a month. Select all answers that apply. – How often do you play the game … ? Proven or confirmed recall – What was the ad/program/game about? Recognition – (Show cover/masthead/article/ad): Have you seen this before?
Different measurement units – How often do you watch television programs that contain violence? [frequency] – And on the days that you watch television programs that contain violence, how much time do you spend on this per day? [duration] – How many anti-smoking advertisements have you seen in the past week? [volume] – How much attention do you pay to news on TV/newspaper articles about national politics? [attention] – List measures
Different reference periods – How many days in a typical (also average, general, or usual) week do you watch the news on TV? versus – How many days in the last week did you watch the news on TV? versus – Do you watch the news on TV at this moment/in the previous hour? [in situ/momentary]
Different response scales – How often do you play video games in an average week? (1) never or almost never, (2) less than once a week, (3) once or twice a week (4) three or four times a week, (5) almost every day or daily [numbered] versus – (1) often, (2) regularly, (3) sometimes, (4) rarely, (5) never [verbal]
Instructions/help – Present vs. missing – Type
Linkage – Combination with content analysis

Slater (2004) summarized a dilemma: "More important, people who process messages with relatively little attention are likely not to remember them in the context of a free recall task, but are more likely to recognize them (Shapiro, 1994a). Therefore, recognition measures are probably less confounded with variables related to attention such as prior interest in the topic than are recall measures, and therefore are closer to the conceptual definition of exposure. One of the principal problems of recognition measures, however, is the tendency of people to report recognizing messages that they in fact have never seen (Shapiro, 1994a)" (p. 170). The use of foils or ringers is a way of estimating the size of this problem (Niederdeppe, 2014; Slater, 2004; Sullivan et al., 2021).

False recognition (a person recognizes media content to which they were not exposed to) is demonstrated in several studies. Measurement error and social desirability can account for this phenomenon, but false recognition can also be real. It is shown, for example, that age predicts recognition error, probably because older people focus on the "meaning of the experience rather than specific attribute details" (Southwell & Langteau, 2008, p. 103).

Measurement unit

The second dimension on which self-reported exposure measures differ is the unit of measurement, which ranges from frequency (how often; how many days), duration (time spent), volume (how much), to attention (degree). It is shown that frequency questions are easier to answer for respondents than duration questions (e.g., Ernala et al., 2020), but they are less informative: people who attend to TV with the same frequency (e.g., every day) may spend considerably different number of hours on the medium.

Reference period

The choice of a reference period (the period for which a respondent is asked an exposure estimate) comes with several issues.

Issue 1: retrospective or momentary (*in situ*)*?*

The first issue refers to the time between media exposure and the exposure question which can be retrospective or momentary. Retrospective questions ask about a respondent's media behavior in the past, say yesterday, last week, or last month (also called "recency" measures), or in a typical/average week (also called "frequency"[1] measures). Retrospective measures have memory issues. Momentary or in situ measures ask for respondent's media exposure at that moment. An advantage is that the time between exposure and recall is as short as possible which minimalizes memory issues and thus may contributes to the validity and reliability of the measure. We come back to momentary measures in the section 4.2.

Issue 2: last versus average/standard

When it comes to the question whether asking about a specific (recent) time period (e.g., last week) is better than asking about typical behavior (e.g., an average week) the evidence is mixed (Juster et al., 2003; Tourangeau et al., 2000). Araujo et al. (2017): "On the one hand, asking about a specific time period in the recent past potentially reduces the cognitive load on respondents, and brings more accurate responses. This recency effect has been found, for example, for estimation of TV exposure (Wonneberger & Irazoqui, 2017). On the other hand, earlier research also found that respondents tend to over-report typical behavior and may actually under-report their behavior when being asked about more recent time periods indicating that asking about a typical week may have higher predictive validity for other outcomes, such as current events knowledge (Althaus & Tewksbury, 2007; Chang & Krosnick, 2003). An alternative explanation for the higher predictive validity of 'typical week' questions could be that this question is confounded with attitudes, such as political interest or involvement, and therefore more strongly correlated with knowledge (Prior, 2009a)" (p. 175).

Price (1993) compared past week and typical week questions (number of days watching television news / reading a newspaper) and concluded that past week questions (being more specific and more recent) generated lower overall reports, and greater variability of media usage. In his study the predictive validity of both types of questions was comparable, and he concluded that the potential atypicality of the narrower time period is "not a serious concern" (p. 615). He also concluded that adding the question "Was this a typical week for you with respect to how often you … ," does not "prove worth the cost of additional interviewing time" (p. 624).

Issue 3: length of the reference period

A third issue related to reference period is its length which can be short (last hour) or long (last year). "Many of [these] studies show a net decrease

1 Note that here "frequency" has another meaning than in the distinction between frequency, duration and volume made above.

in the number of reported events per unit time as the length of the response period increases, probably due to forgetting (see Sudman & Bradburn, 1973, for a meta-analysis). As Neter and Waksberg (1964) note, longer references periods increase not only the amount of time over which respondents must remember events but also the total number of events they must recall, and both variables are likely to reduce accuracy" (Tourangeau et al., 2000, p. 86). Studies have found that the accuracy of self-reports of frequent, mundane behaviors is facilitated when shorter (and recent) reference periods are used (Blair & Burton, 1987; Burton & Blair, 1991; Converse & Presser, 1986; Price, 1993; Schwarz, 2007; Schwarz & Oyserman, 2001).

Issue 4: decomposition of reference period
It is shown that accuracy of recall is facilitated if a reference period is divided into meaningful components — for instance weekdays and weekend — because this facilitates the retrieval of all relevant instances (Belli et al., 2000; Tourangeau et al., 2000; Vanden Abelee et al., 2013). On the other hand, the decomposition of media exposure questions into sub questions may inflate people's self-reports. "Decomposition is a questionnaire design strategy, often advocated in survey research, in which behavioural frequency reports for a category are broken down by asking about the behavioural frequencies for subcategories .... This strategy reliably results in higher-frequency reports, i.e., the sum of the events reported in response to the subcategory questions exceeds the number of events reported in response to the general question" (Belli et al., 2000, p. 295). In the study of Vanden Abeele et al. (2013) into respondents' telephone calls, decomposed questions increased over-reporting bias relative to undecomposed questions. See also: Bradburn et al. (1987), Prior (2013), Schwarz and Oyserman (2001), and Tversky and Koehler's (1994) unpacking effect.

Content specificity

We continue with the next dimension on which media exposure questions can be distinguished: content specificity. Conceptually, media exposure may relate to different types and levels of content (see Chapter 2). With respect to measurement: global measures such as "How often do you watch television?" have the disadvantages of lack of specificity, lack of control for third variables, and reverse causality (e.g., Slater, 2014; Valkenburg & Peter, 2013a). More specific measures (programs, titles, specific media, genres) provide more information about the content to which a media user is exposed, also because it allows for a content analysis which can be combined with the exposure information. Romantan et al. (2008) found that specific questions may be most useful considering criterion validity, face validity, survey costs, and respondent burden.

Disadvantages also apply. In longitudinal studies, a problem may be that the content of the program, the printed title, the game, or the social media platform changes. It is also possible that social desirability plays a role when respondents are asked about their exposure to socially (un)desirable content. Another issue related to more specific exposure measures is mentioned by Fikkers et al. (2015): Respondents "not only have to recall when and how long they were playing games or watching television but also more specifically which kind of content they were consuming at those times, and only report those instances in which content was violent. This brings an additional cognitive task to the answering process that may affect the reliability and validity of the resulting answers" (p. 120).

Opree et al. (2021) investigated several recall measures for children's exposure to television and internet for three levels of content specificity: medium (e.g., TV, internet), broad content (e.g., channels, programs, websites), and specific content (e.g., commercial channels, websites with the most advertisements). They evaluated these measures on the basis of several criteria including test-retest reliability, and content, criterion and construct validity. Their panel study (data collected from 165 children 8–11 years of age) showed that all measures provided solid estimates for children's television and internet advertising exposure. The authors conclude that *television advertising* exposure can best be measured by asking children how often they watch certain popular (commercial) network stations (either weighting or not their scores for advertising density) rather than programs because broadcasting schedules change, and because "the more content-specific measures did not really outperform the broad measures in terms of reliability and validity" (p. 670). Exposure *to internet advertising* can best be measured by asking children how often they use the internet or asking them how often they visit certain popular websites, weighting their scores for advertising density.

Content specificity: list measures

Dilliplane et al. (2013) presented respondents with a list of TV programs and asked them to indicate whether they had watched each of the programs "regularly" (at least once a month) and named this method the "program list technique." The procedure has been adopted by the National Annenberg Election Survey (NAES) in its 2008 online panel (Guess, 2015) and the American National Election Study (ANES) for its 2012 election survey (Prior, 2013). The strong points of this approach are that respondents do not have to decide for themselves which programs include for instance "news," and that it might be easier for respondents to recall programs than day times. See Table 4.2 for the questions.

Based on data from the 2008 National Annenberg Election Study (NAES), a five-wave panel, the authors showed that the program list method has strong true-score reliability (.88 on average), along with good content validity, predictive validity (criterion: change in political knowledge), and discriminant validity (Dilliplane et al., 2013). See for a similar approach: Greenberg et al. (1968) and Avery et al. (2012).

Table 4.2 The program list method

"Respondents were first asked, "From which of the following sources have you heard anything about the presidential campaign?" Respondents checked all of the categories (presented in randomized order) that applied to them: (1) television news programs (morning or evening); (2) newspapers; (3) television talk shows, public affairs or news analysis programs; (4) Internet sites, chat rooms, or blogs; (5) radio news or radio talk shows; (6) news magazines; and (7) have not heard anything about the presidential campaign. Respondents who selected categories (1) or (3) were then shown a screen with a list of TV programs asking, "Which of the following programs do you watch regularly on television? Please check any that you watch at least once a month." Later screens were sprinkled throughout the survey in random order to avoid boredom from checking off programs from multiple screens in a row."

Source: Dilliplane et al. (2013, p. 238).

Prior (2013) criticized the list method. He argues that the method has a number of aspects that are difficult for respondents: (1) they have to understand what content belongs to "presidential campaign," (2) have to remember the name of the program that they were watching, (3) and have to remember to have watched the program in a specific time period. He also points to the fact that the number of programs is not a good estimator of the extent to which a person is exposed to campaign information (e.g., because some will watch the program daily, other seldomly). In other words: the number of programs is an indication of the breadth rather than the amount of the media consumption.

Andersen et al. (2016) refined the list technique in three ways: (1) by including other media types than television (newspapers, websites, and radio), (2) by including a time period ("last week"), and (3) measuring the amount/frequency of exposure ("how many days"). According to the authors their "frequency list measure" has the advantages of measuring different media outlets, a reduced time period between exposure and measurement (last week), the possibility to connect exposure with a content analysis of the past week media content, the possibility to track exposure changes over time (e.g., in a campaign), and the possibility to take intensity (frequency) into account.

Their study showed that their version did not take more time than the original list technique, that the aggregated level of exposure to different sources were not substantially different from the original list technique, and that the predictive validity of the refined technique was higher. It is clear that the selection of programs (necessarily limited which is problematic because of the increasingly fragmented audiences) that are presented to the respondents is a crucial decision when using this method, that social media are a challenge, and that the degree of attention is not taken into account. Also under this method, respondents tend to overreport their media use.

Fikkers et al. (2015) mentioned two other options for selecting specific media content for exposure questions: (a) ask respondents for their favorites (see also Anderson & Dill, 2000), and (b) select the most popular programs or games as reported by media rating systems. They tested these two options and a direct estimate (television programs / games that contain violence) in a study (n = 238 early adolescents) and concluded, “For game violence, the three self-report measures were reliable and valid. For television violence, only direct estimates achieved test-retest reliability and construct validity” (p. 117).

Instead of presenting respondents with a list of programs, open-ended questions are also possible as in Guess (2015). In his study, the open-ended question outperformed other survey-based measures of online media exposure. The author argues: “One advantage of this type of question is that since respondents are not presented with a preselected list of choices, it reduces errors in judgment caused by the familiarity of a given item: ‘Although true frequency can increase familiarity, so can factors like ease of perceiving an item, expectation induced by context, and probably many other variables’ (Tourangeau et al., 2000, pp. 142–143)” (p. 62).

Response scale

Answer scales can be numeric (e.g., 3 out of 5 workdays) or verbal (often, regularly, ...). Coromina and Saris (2009) compared these two types (one numeric and two verbal) for media exposure variables in a pilot study of the European Social Survey (see Table 4.3), and concluded that the numerical categories scale was the most, and the verbal categories scale the least reliable and valid.

Another response scale issue: Schwarz (1999) and Schwarz et al. (1985) argue that respondents might interpret the middle of the scale presented to them as the "average" or "usual" category, which can stimulate socially desirable answers.

Other general issues of response scales in self-reports are discussed in Bergkvist and Langner (2020): (1) the use of inappropriate numerical response scales, (2) mixing unipolar and bipolar response scales, (3) the use of antecedent and outcome items, and (4) the inconsistent use of response scale endpoint qualifiers. DeCastellarnau (2018) provides the main conclusions from the literature on the impact of response scale characteristics on data quality.

Table 4.3 Three response scales in the study of Coromina and Saris (2009)

Numerical categories scale	**Numerical open question** Write in hours [ ] and minutes [ ]	**Verbal categories**
– No time at all	… hour and … minutes	– No time at all
– Less than 1/2 h		– Very little time
– 1/2 to 1 h		– A little time
– 1 to 1½ h		– Some time
– 1½ to 2 h		– Quite a lot of time
– 2h to 2½ h		– A lot of time
– 2½ to 3 h		– A great deal of time
– More than 3 h		

Source: Coromina and Saris (2009).

Instructions/help

Self-reports of media exposure differ in the way instructions or help are given. Aided recall measures in which respondents are provided with a description of a message are an example of helping respondents recollecting their exposure. Other possibilities include the recommendations by Belli (1998), Burton and Blair (1991), and Schwarz et al. (1985) who have suggested that the quality of self-reports may be increased by providing an anchor that gives information about other people's behavior, for instance by offering population averages.

Another option is presenting respondents contextual cues by asking questions activating memories of past behavior (e.g., Araujo et al., 2017; Jerit et al., 2016; Menon & Yorkston, 2000; Neter & Waksberg, 1964; Potts & Seger, 2013; Sudman et al., 1984). Schwarz and Oyserman (2001) suggested cues such as "what happened, where it happened, and who was involved"

(p. 138). Kahneman et al. (2004) developed the Day Construction Method in which respondents were asked to construct a short diary of the previous day: "Think of your day as a continuous series of scenes or episodes in a film. Give each episode a brief name that will help you remember it (for example, 'commuting to work', or 'at lunch with B'). Write down the approximate times at which each episode began and ended. The episodes people identify usually last between 15 minutes and 2 hours. Indications of the end of an episode might be going to a different location, ending one activity and starting another, or a change in the people you are interacting with" (p. 1777).

In the experiment of Griffioen et al. (2020a) subjects were videotaped (a camera installed right above the participant's seat), and this video footage of their actions (in combination with in-app logs) was later used to interview them to help them answer questions about their phone and social media use during the waiting period.

Media diary

A media diary is a tool in which respondents daily report their media use of that day or the previous day during a certain period of time (for example, a week). A typical design of a media diary consists of a grid with rows indicating channels, programs, newspaper titles, magazine titles, or any other media unit, and columns indicating time (periods) of the day (e.g., 15-minute intervals) (Juster & Stafford, 1991; Slater, 2004).

Because the time between exposure and response is short (the same or previous day), recall issues are less problematic than in surveys. Fikkers et al. (2015): "The strength of media diaries lies in two elements that are known to improve recall. First, media diaries capitalize on the autobiographical structure of our memory. By encouraging participants to think about their day, a rich network of associations is activated, which increases the likelihood that individual episodes of media use are retrieved (Schwarz & Oyserman, 2001). Second, because media diaries tend to be filled out on the day itself or the day after, this short and recent reference period improves the likelihood of accurate recall (Schwarz & Oyserman, 2001)" (p. 122).

Diaries have the advantage of not disrupting normal activities and covering a full day instead of sampled moments, but have possibly more recall issues (Kahneman et al., 2004). Coding of the diary can be time consuming for the researcher.

Different anchors

Prior (2009a) studied how researchers can help respondents by providing population frequencies and encouraging comparison with others (see

Table 4.4). His findings showed that only providing a combination of both types of information somewhat helps against overestimation.

Araujo et al. (2017) applied three anchors in their study: "Before answering the main questions, a random selection of 50% of the respondents first answered three anchoring questions. The first two questions asked the respondent to think about the previous day/week, and indicate what type of day/week it was (e.g., normal day/week, holidays, being sick, etc.). In addition, respondents were asked to think back in which situations they made use of their private computer or tablet (e.g., at home, at work, while commuting, etc.). This anchoring procedure aims at improving recall and has been used successfully applied in earlier studies (Bronner & Neijens, 2006)" (p. 177). In the study of Araujo et al. (2017), however, the anchoring hypothesis was not supported.

Table 4.4 Anchors provided by Prior (2009a)

Anchor 1: Television news audiences have declined a lot lately. Few Americans watch the national network news on a typical weekday evening. Anchor 2: Television news audiences have declined a lot lately. Less than one out of every ten Americans watches the national network news on a typical weekday evening. Anchor 3: With all that's going on in the world these days, many Americans watch the national network news on a typical weekday evening. Anchor 4: Just the first sentence of Anchor 2. Anchor 5: … "large audiences" instead of "many Americans" (see Anchor 3). Anchor 6: With all that's going on in people's lives these days, some watch the national network news on a typical evening, while others don't.
Findings: Only the combination of population frequency and explicit reference to other people remedies [anchors 1–3] flawed estimation of news exposure.

Source: Prior (2009a).

Linkage: combining self-reports with media content

Some scholars argue that it is a good idea to combine media exposure measures with the content of the media a person is exposed to as this allows for a more accurate and specific analysis of the use and impact of media. De Vreese et al.'s (2017) advices for the so-called linkage of exposure data with content analysis data are "(1) include exposure measures in the survey in the most granular and detailed way possible (time and resources permitting) to allow for multiple approaches in subsequent linkages; (2) focus on substantive content features" [such as 'visibility of actors', 'tone', 'frame', etcetera] "first in any linkage analysis and add formal features" [such as prominence, size/length, or position] "only later in the analysis" (p. 240). Theory should be the guiding principle here.

For the combination of both types of data, de Vreese et al. (2017) recommend taking theoretical considerations as a starting point: "If the endeavor is to test for the effects of content in a specific medium (e.g., TV news vs. newspapers), it makes sense to organize the exposure measures per medium (e.g., Andersen et al., 2016). If the interest is to test for the effects of content in a specific genre, it may make sense to distinguish between quality outlets (e.g., public service news and broadsheet papers vs. commercial news and tabloids). If the interest is to test for the effects of content exposure all together, it may make sense to create a single exposure measure. The same is true when there are no significant content differences across outlets (see e.g., Schuck & de Vreese, 2008)" (p. 226). Ethical issues of data linking may include the lack of informed consent of respondents. Measurement error issues in linkage analysis are discussed in Bachl and Scharkow (2019). See Otto et al. (2022) for linkage analysis where respondents upload (screen) shots. Otto et al. (2023) focus on linking different sources, practical challenges, and design decisions, as well as analytical opportunities and complexities of linked datasets of digital trace data and surveys.

4.2 Momentary or in situ measurement of media exposure

Momentary or in situ measurement of media exposure is seen as an ideally suited self-report method for measuring media exposure in the current media landscape, which is characterized by superficial and short-term media contacts amidst an enormous range of media. An early example of in situ — or momentary — methods are the telephone coincidentals in the USA in the 1930s which asked people what they were listening to at the time of the call (Webster et al., 2014). Although not representative for an individual's media behavior, given randomly sampled time intervals and respondents, coincidentals produce adequate aggregate statistics.

Experience Sampling Method / Ecological Momentary Assessment

Today, the Experience Sampling Method (ESM) (Csikszentmihalyi & Larson, 1987; Kubey et al., 1996; Link et al., 2014) is a popular method for momentary measurement. ESM can be defined as "a research procedure that consists of asking individuals to provide systematic self-reports at random occasions during the waking hours of a normal week" (Larson & Csikszentmihalyi, 1983, p. 41), or "a method of data collection in which respondents repeatedly report on behaviour, cognitions, and emotions over a certain period of time across several situations. Each time they are alerted, subjects are asked to answer a short questionnaire (called Experience Sampling Form, ESF) with as little delay as possible. Hence, this approach samples situations from users' everyday lives as data are collected in natural settings and across situations" (Naab et al., 2019, p. 129). Other terms for these approaches: "Ecological Momentary Assessment" (EMA, e.g., Shiffman et al., 2008; Stone & Shiffman 1994), or "Real-Time Response Measurement" (RTR, e.g., Biocca et al., 1994).

Potential issues with momentary methods include its dependence on representative sampling of situations (with associated questions as: how many weeks, and how often during the day should the respondent be prompted to submit their media behavior), its disruptive character, and compliance issues because intensive repeated measurement places high demands on participants (Naab et al., 2019; Scollon et al., 2003; Stone et al., 2007). To keep the demand on respondents as low as possible, the concepts in ESM are often measured with one item (Fisher & To, 2012; Gabriel et al., 2019; van Hooff et al., 2007). Validation of single-item measures for ESM is central to the project of Wolfers and Baumgartner (2023).

In their study of the use of mobile devices as self-report data collections tools in repeated measurement designs, Schnauber-Stockmann and Karnowski (2020) reviewed the media and communication literature and found 31 studies that applied these methods, a number that was already three times as high in 2022 (n = 101) (Schnauber-Stockmann et al., manuscript submitted for publication). Based on their 2020 research, the authors developed a typology of these designs, distinguishing between "mobile diary," "mobile experience sampling," and "mobile real-time response," each with its own characteristics, such as number of study days, number of entries per day, type of prompts, sampling, and coverage. Their analysis showed considerable variations in the application of these methods (see Table 4.5). The smartphone was the most applied device.

Table 4.5 Variations in applications of mobile data collection methods for self-reports of media exposure

Aspect	Variations found
Field period	From less than one day up to 42 days
Sampling of moments	Fixed, random, event-contingent
Prompts	Participant-initiated vs. researcher-initiated
	Time-contingent vs. event-contingent
	From 1 per day up to 48 per day
Reference object / coverage / unit of measurement	Whole day, current, event

Source: Schnauber-Stockmann and Karnowski (2020).

Think-aloud

Think-aloud procedures ask respondents to verbalize the thoughts they experience(d) while performing a task (Ericsson & Simon, 1993; van Someren et al., 1994). In these procedures it is important "that subjects are not asked to justify or explain their thoughts or way of thinking. Thus, it keeps rationalizations by the subject to a minimum. Of equal importance is the fact that this technique is as non-obtrusive as is possible. The only probe subjects receive is the instruction to talk aloud" (Schaap, 2004, p. 120). The "protocols" that result from think-aloud procedures are transcribed, coded, and analyzed. Think-aloud procedures can be combined with successive interviews for a more precise analysis of verbal protocols (Branch, 2000; Hoppmann, 2009).

Ericsson and Simon (1993) distinguish two types of think-aloud procedures: concurrent (the cognitive processes are verbalized directly, i.e., at the same time as performing the task) and retrospective (participants verbalize their thoughts after completion of the task) as in the "thought-listing technique" (e.g., Valkenburg et al., 1999). The former has the advantage that it does not create memory issues, but might interfere with the task at hand. Retrospective think-aloud has the advantage of being less intrusive, but may suffer from memory problems.

An illustration of the think-aloud methodology can be found in Table 4.6 that summarizes how the method is used in a study by Eveland and Dunwoody (2000). It also shows that think-aloud data can miss the intended target.

Table 4.6 Think-aloud study

The authors examined information processing on the internet. A group of high and low web users, males and females (*n* = 15) were invited to the lab for an individual session of 90 minutes. First participants engaged in several practice tasks to familiarize themselves with the process of thinking aloud. The main task "placed participants on the home page of The Why Files Web site, but participants were informed that they were free to navigate from there to anywhere on the Web. The task lasted about 30 minutes for most participants" (p. 228). An audiotape recording and a video recording of facial expressions and images on the screen were made. The transcripts of the think-aloud procedure were segmented into "'thought' units" (a sentence, a clause of a sentence, or a phrase). Two coders independently unitized all think-aloud comments and several information processing variables (e.g., orientation, elaboration) were coded for each thought unit. "Most of the thoughts generated by the think aloud procedure referred to the content of the sites instead of their structure. Only about 2.5% of comments pertained to the Web generally, the browser software, or the computer hardware" (p. 233).

Source: Eveland and Dunwoody (2000).

Pros of the think-aloud technique are that respondents can verbalize their thoughts in their own language, that it is a concurrent procedure (no memory issues), and that the data can be analyzed quantitatively and qualitatively. There are also possible cons (Benbunan-Fich, 2001; Branch, 2000; Hoppman, 2009; van Someren et al., 1994). First the interference issue mentioned above: the think-aloud task might affect "the ongoing cognitive process thereby potentially changing the outcome of that process" (Rozendaal et al., 2012, p. 203; see also Shapiro, 1994b). Second, some thoughts are non-verbal and complicated to express. Third, verbalization is a slower process than the underlying cognitive processes. Fourth, think-aloud makes information processing more conscious, which implies a problem for measuring habits as those tend to be processed peripherally (Woelke & Pelzer, 2020). Five, the sample size in think-aloud studies is usually limited (between 10 and 30 respondents), which is less of a problem in some type of studies (for instance, usability studies, Nielsen, 1994), but offer a generalization problem in media exposure studies. Six, social desirability may be a problem as well. Seven, a complication for the analysis is that the data is rather extensive and ill-structured.

4.3 Quality issues of self-reports

Conceptual, empirical, and practical issues affect the quality of self-reports (Niederdeppe, 2014).

Conceptual issues

Let's start with Slater who addresses a crucial dilemma. The media exposure definition ("the extent to which audience members have encountered specific messages or classes of messages/media content") suggests that "exposure refers to a person's merely encountering the messages, whether or not they are noticed enough to be remembered. After all, noticing the relevant messages in the communication environment is almost certainly confounded with variables that may predict attention to the content of that message, such as prior knowledge or involvement with the topic. It is also quite possible that exposure may leave an affective if not a cognitive impression of some kind, even if the messages have not been attended too well enough to be remembered. However, if messages are not processed thoroughly enough to be recalled, how can exposure be self-reported?" (Slater, 2004, pp. 168–169). In other words: on the one hand it is not possible to measure levels of exposure that have not been attended well enough to be remembered, and on the other, levels of exposure that are remembered are most probably related to interest and involvement with the topic, making the relationship between media exposure and media effects partly spurious because it is confounded with the effect of interest in the topic.

A related issue is that exposure in media effect models "is typically a mediating, rather than a truly exogenous, variable: It may be substantially influenced by both baseline scores on an outcome variable and by many possible third variables (also known as endogeneity)" (Slater, 2004, p. 171; see also Slater, 2014).

Empirical issues

In addition to conceptual issues, there are empirical issues. Studies into the accuracy, validity and reliability of self-reports demonstrate that these measures overestimate media exposure, show moderate reliability, and that alternative media exposure measures correlate only moderately (e.g., Araujo et al., 2017; Bechtel et al., 1972; Boase & Ling, 2013; Junco, 2013; Lee et al., 2008; Niederdeppe, 2014; Price, 1993; Prior, 2009b; Scharkow, 2016; Vanden Abeele et al., 2013; van der Voort & Vooijs, 1990; Verbeij et al., 2021). Let's have a look at the details.

Problems with self-reporting are much like problems answering survey questions about frequency of past behavior. According to survey answering models (e.g., Cannell et al., 1981; Schwarz & Oyserman, 2001; Tourangeau et al., 2000), respondents have to (1) understand the question, (2) recall the relevant behavior, (3) estimate the frequency of the relevant behavior, (4) map the frequency onto the response alternatives, and (5) report either their candid answer or a socially desirable answer. Problems may arise during each of these activities. We will discuss them successively below.

Understanding the question. First, people must understand what the researcher means with for example the news, or violence, and which programs contain (sufficient) violence to justify inclusion. This is problematic, because media content referred to in typical media exposure questions is often too broad to allow unambiguous interpretation. "Watching TV news," for instance, is a broad category which implies that people with the same score may be exposed to very different media content (see also Belson, 1981). A study by Vraga et al. (2016b) found that the lines between different types of content may not be clear to respondents. They asked their (student) sample to classify Facebook posts and found that "users and researchers often agree on defining social and political content but are more likely to disagree on categorizing news content" (see also Mutz & Young, 2011). Another source of confusion arises if other terms in self-report questions are unclear (for instance: what is a typical/regular/average week or what counts as watching?) or vague (when researchers use the categories "seldom," "regularly," or "often").

These problems are illustrated in studies of Belson (1981, 1986), who interviewed respondents after they had filled in a questionnaire with standard media exposure questions and found a substantial number of different and incorrect interpretations of the questions which made it problematic to interpret and compare their responses in an univocal way.

Recall and estimation. When respondents are asked to report the frequency of incidents (e.g., media exposure in a typical week), they may apply several strategies, such as recall of specific information and count, estimation based on generic information (such as "two visits to the dentist are recommended, thus the answer is 2"), tally (parents usually know their number of children without counting them), and extrapolation (Tourangeau et al., 2000, pp. 146–150). Bradburn et al. (1987): "In many situations of importance in survey work, respondents are simply unable to retrieve and count separate incidents. Instead, they use the fragmentary information that they have and extrapolate as necessary" … which is "at best inexact and at worst misleading" (p. 161). Research on the perception of duration (e.g., how many hours in a typical day you play games?) shows that pleasant, numerous, variable, and difficult activities are perceived shorter than unpleasant, few, monotonous, and easy stimuli (Galinat & Borg, 1987). Table 4.7 shows these and other characteristics of the stimulus, study, and receiver which have been found to distort adequate recall and estimation of frequency and duration of (media) events.

Table 4.7 Examples of factors that influence accuracy of recall of media exposure

Stimulus
· dramatic impact, saliency, distinctness of events
· amorphous events
· irregular events
· modality (e.g., auditory stimuli may appear to last longer than visual stimuli)
· availability (respondent's judgement that events that are easy to remember must be frequent)
· pleasant/enjoyment (time passes faster)
· prestige of the medium
· underestimating of high frequency events, overestimating of low frequency events
· time interval (shorter intervals tend to be overestimated while longer intervals tend to be underestimated
· recentness (intervals longer ago seem shorter)
· telescoping (displacement of events in time, which can be backwards (too remote) or forwards (too recent)
· publication interval (e.g., weekly, monthly)
Study
· characteristics of exposure question and answer scale (the variables discussed in this chapter)
· response time
· response order (primacy and recency effects)
Receiver
· interest and motivation
· active consumption (associated with less accurate recall because of cognitive limitations)
· habit
· self-presentation concerns
· individual level characteristics such as age, education, income and developmental disorders
· occasional versus frequent readers
· subscribers versus non-subscribers
· type of information processer (memory based vs. on-line)

Sources: Araujo et al. (2017), Bradburn and Sudman (1979), Burton and Blair (1991), Cannell et al. (1981), Clancy et al. (1979), Clay et al. (2013), Corlett (1964), Corlett and Osborne (1966), Deng et al. (2019), Fikkers et al. (2015), Fortin and Rousseau (1998), Hornik (1984), Kaye et al. (2020), Lin et al. (2015), Marder (1967), McGlathery (1967), Naab et al. (2019), Price (1993), Robinson and Clore (2002), Schwarz and Oyserman (2001), Sekely and Blakney (1994), Thompson et al. (1988), Thompson et al. (1996), Tversky and Kahneman (1973), Vanden Abeele et al. (2013), Verbeij et al. (2021), Whipple and McManamon (1992), Wonneberger and Irazoqui (2017).

Not only cognitive and motivational problems cause recall problems, the current media landscape also plays a role. Niederdeppe (2016) illustrates this with a striking example (see Table 4.8).

Table 4.8 Which source?

"Imagine the following scenario: 'A friend posts a link to a usatoday.com article about Donald Trump's immigration policy proposal to your Facebook timeline. You click on the article, read it, and glance at the lengthy set of comments that follow it. The next day you watch a feature on your local TV news about reactions to the proposal on the "Twittersphere" and chat briefly with a colleague at work about Trump's platform more generally. Now suppose a researcher calls you for a telephone survey the next day and asks you (a) what you've heard about Donald Trump's stance on immigration, and (b) where you heard about it. What would you say? There would seem to be at least five credible answers for the source: a newspaper, website, multiple social media platforms, friend/colleague, and local TV news. Now suppose that the researcher asked you about it four weeks later. What is the likelihood that you would remember anything about the details?'"

Source: Niederdeppe (2016, p. 170).

Map onto the response alternatives. A fourth challenge in answering survey questions about media attention arises when the response categories of the self-report questions are vague, such as when researchers use the categories "rarely," "regularly," or "often."

Social desirability and other motivational issues. Fifth and finally, there may be motivational considerations in answering survey questions. Some respondents do not want to report media exposure to specific media content such as low prestige publications, violence or pornography. Other motivational problems arise when people are not willing to think hard to arrive at the correct answer. In media exposure research this maybe a problem as respondents can get exhausted or annoyed with answering long questionnaires about the many media outlets that a person possibly is exposed to and therefore have a tendency to underreport their media behavior, a phenomenon known as "satisficing" (Krosnick, 1991).

Reliability

As indicated in Chapter 3, panel studies make it possible to estimate the test-retest reliability of self-reports. Scharkow (2019) conducted a meta-analysis of 33 panel studies on exposure to television, radio, newspaper, and internet (video games, home videos or DVDs, music, or books were not included in the analysis). He found that self-report media exposure measures were moderately reliable (Heise reliability coefficient = 0.69) and highly stable (rank order stability = .90), "supporting the overall conclusion that media exposure is a very stable behaviour, even in a high-choice media environment" (p. 206). These findings were very similar to those of Lee et al. (2008). Scharkow also found that reliability was:

- higher for adult samples than for adolescents,
- higher for general media use than for specific outlets (in adult samples),
- higher for television and internet than print and radio (for adolescents),
- not influenced by response format (days per week vs. open-ended vs. ordinal frequency scales) (for adults),
- less for open ended questions (hours of use in a given period) (for adolescents).

Validity

Studies into the validity of self-report measures have a long history. Already fifty years ago, Bechtel et al. (1972) compared self-report diaries of daily viewing habits with a video of the behavior in the home of 20 families in Missouri for six days. The equipment operator monitored the recording from a rented truck parked in the driveway behind or beside the house. The study noted that self-reports (diaries) overestimated actual time viewing by about 25%. Prior (2009b) compared self-reported exposure measures with Nielson audience ratings (aggregate level study) and also found that survey estimates of network news exposure in the USA "exaggerate exposure by a factor of 3 on average and as much as eightfold for some demographics" (p. 130).[2]

In the last decade, the number of studies that made comparisons between self-reports and other methods has exploded. We give four examples and then we present a meta-analysis and a MTMM analysis. First, Wonneberger and Irazoqui (2017) compared self-reports and people meter measures of TV exposure (frequency and duration) at the individual level. They found a clear tendency to overreport frequency of watching and underreport viewing duration, in line with the higher cognitive demand of the duration question. Response errors were systematically related to respondent characteristics (age, education, and income) and showed a social desirability effect. Those with the most stable viewing patterns were the most accurate. Heavy viewers underreport their viewing behavior; light viewers overreport.

2 Another illustration of overreporting are studies that include "false" messages that were not part of a campaign or media content. Niederdeppe (2014): "Typically, between 7% and 16% of respondents falsely report having seen a false message (e.g., Brown et al., 1990; Southwell et al., 2002; Thrasher et al., 2011)" (p. 148).

Second, Boase and Ling (2013) made a comparison between self-report measures and server log data from a telephone company. The authors compared these measures for outgoing calls and outgoing text messages. They concluded that self-reports were only moderately correlated with actual behavior, varied more widely than actual behavior, and were prone to overreporting.

Third, Wenz et al. (2024) evaluated how passively collected smartphone usage data (captured by a tracking app) compared to self-reported measures of smartphone use. They found that amount of use is considerably overreported in the survey data and that alignment between the two measures varies by type of activity (such as making phone calls, messaging, visiting websites, or shopping). The results also showed that the fit between the measures is systematically related to individuals' sociodemographic characteristics, including age, gender, and education.

Fourth, Verbeij et al. (2022) compared the predictive validity of self-reported and digital trace measures of time spent with social media (n = 159 adolescents). As (presumed) effect variables they included self-esteem, well-being, and friendship closeness. "Using an n = 1 method of analysis, we investigated the correspondence on a between-person, within-person, and person-specific level. Although our results confirmed the poor convergent validity of self-reported TSM [time spent with social media] reported earlier, we found that self-reports of TSM had comparable predictive validity to digital trace measures on all three levels" (p. 1).

Meta-analysis

Studies that made comparisons between self-reports and digital trace data included different media (e.g., the internet, television, smartphone, social media), type of content, platforms, populations (e.g., general population, adolescents, students), validity criteria, type of measure (e.g., frequency, duration; retrospective, momentary), frequency of behavior (light, heavy), demographics (age, gender, education), and were conducted in different time periods.

The general conclusions of these studies are that self-report data on digital media exposure is (much) higher than the corresponding digital trace data, has lower variation, and that these two types of data only correlate moderately (e.g., Araujo et al., 2017; Boase & Ling, 2013; Burnell et al., 2021; Cohen & Lemish, 2003; Deng et al., 2019; Ernala et al., 2020; Haenschen, 2020; Johannes et al., 2021; Jurgens et al., 2019; Kobayashi & Boase, 2012; Naab et al., 2019; Parry et al., 2021; Parslow et al., 2003; Scharkow, 2016; Sewall et al., 2020; Song & Cho, 2021; Timotijevic et al., 2009; Vanden Abeele et al., 2013; Verbeij et al., 2021, 2022; Vraga et al., 2016a). Ohme et al. (2021), however, found that people mostly underreport the duration of usage and frequency of incidents such as checking their phone or receiving push messages, which the authors tentatively explain by the higher ubiquity of smartphones compared to 10 years earlier.

Parry et al. (2021) conducted a meta-analysis of studies that made comparisons between self-reported and log-based measures of *social media use* obtained via digital trace data collection methods. Their study shows that self-reports are only moderately correlated with digital trace data. Because their correlation data showed a high level of heterogeneity, they tested various moderators (possible explanations for the different correlations between the two types of measures), but none of them were significant (party due to a lack of (sufficient) cases). Table 4.9 shows the details.

MultiTrait MultiMethod (MTTM) analysis

Cernat et al. (2024) estimated the measurement quality of survey and digital trace data on smartphone usage with a MultiTrait MultiMethod (MTMM) model: "The experimental design included five topics relating to the use of smartphones (traits) measured with five methods: three different survey scales (a 5- and a 7-point frequency scale and an open-ended question on duration) and two measures from digital trace data (frequency and duration). We show that surveys and digital trace data measures have very low correlation with each other. We also show that all measures are far from perfect and, while digital trace data appears to have often better quality compared to surveys, that is not always the case" (p. 1).

Momentary methods: compliance and validity

Below we discuss quality issues of momentary methods such ESM/EMA and think-aloud.

Compliance

Due to the intrusiveness of intensive repeated measurement as applied in momentary methods, compliance with the research tasks (see Chapter 3) is a potential quality problem. Rintala et al. (2019) conducted a meta-analysis into this issue based on 10 studies with a total of 92,394 momentary assessments from 1717 participants with different mental health conditions. They defined compliance as having a recorded response time that fell within a time window of 5 min. before and 15 min. after the beep. Their findings showed "acceptable compliance [overall compliance 78%] in an ESM protocol of 4 to 6 study days with a high frequency of 10 assessments per day despite

Table 4.9 Comparing self-reported and log-based measures of social media use (meta-analysis by Parry et al., 2021)

Data

Automated search on five broad bibliographic databases: PubMed, Scopus, PsychInfo, Communication & Mass Media Complete, and the ACM Digital Library. To target unpublished (grey) literature they used the ProQuest Dissertations & Theses A&I database. This resulted in 66 effect sizes from 44 studies (the first study they found was published in 2007, the last in 2020).

Moderators

(within brackets number of studies)

- Unit: duration 47; volume 19
- Medium: phone 49; social media 13, internet 2, game 1, computer 1
- Self-report form: single estimate 60, scale 6
- Study population (general 4, adult 38, student 15, adolescents 2, unknown 7)
- Data collection (data donation 16, direct tracking 30, supplied data 20)
- Logging method (built in tool 16, custom built tool 15, operator or platform data 20, third party tool 14, other 1)

Findings

1. Self-reported media use correlates with logged measurements only moderately ($r = .38$).
2. Measures of problematic media use (e.g., excessive use or other conceptions of problematic use) show an even smaller association with usage logs ($r = .25$).
3. Accuracy: about 6% of self-reports are within a margin of error of 5% of the equivalent logged value, indicating that, when asked to estimate their usage, participants are rarely accurate.
4. There is a high level of heterogeneity for the correlation between self-reported and logged media use and for the accuracy of the scores.
5. Moderation analyses did not show significant results. The small number of cases for some moderator categories may partly explain the lack of moderator effects.

Source: Parry et al. (2021). The list of References in their article shows the original studies.

fluctuations across and within study days" (p. 226). Compliance declined across days and varied depending on the time within a day. Females and older participants were slightly more compliant.

In a related article, based on seven studies with a total of 72,954 ESM/EMA observations from 1,354 participants, the authors reported 86% overall compliance (to beeps where a subject was compliant at the previous beep). Their findings suggest that disruption of the beep, being away from home, medication use, and inter-prompt interval may reduce the likelihood of compliance to the subsequent beep (Rintala et al., 2020).

Van Roekel et al. (2019) examined 23 momentary studies among adolescents published in 2017. Compliance rates varied between 51.6% and 92.0%

(M = 74.0%) which is "similar to what has been found in adult samples (Hufford et al., 2001)" (p. 566).

Wrzus and Neubauer (2023) also conducted a meta-analysis (including k = 477 articles; 496 samples, total *n* = 677,536). The results showed that on average EMA studies scheduled six assessments per day, lasted for seven days, and obtained a compliance of 79%. Compliance was significantly higher in studies providing financial incentives. The number of assessments did not predict compliance or dropout rates.

Intensive compliance-enhancing techniques were successfully applied in the study by Bij de Vaate et al. (2023) in which adolescents received one questionnaire per day for 100 consecutive days. Participants completed 30,802 daily diary questionnaires during the 100 days, resulting in a compliance rate of 83.1%. Non-compliance was partly due to human factors, such as being ill, and technical factors, such as lack of internet access. Compliance was monitored daily. "Members of the research team were available via WhatsApp, telephone, and e-mail to answer any questions or problems. To increase compliance, the investigators regularly contacted participants with motivational messages (e.g., their own weekly response rate), and contacted participants, for example, if they had missed three consecutive questionnaires to inquire about technical issues with the application" (p. 12). Incentives included €1 for each completed daily diary questionnaire, various other bonuses, and lotteries. Careful and extensive intake of participants may also have contributed to the high compliance rate (Bülow et al., 2024).

Validity

Two recent studies compared retrospective and momentary self-reports of mobile social media use. Naab et al. (2019) studied these measures for Facebook, WhatsApp, and YouTube and found low agreement. "Overall, we observed a consistent pattern of higher estimates in retrospect as compared to individual averages of in-situ reports. The absolute magnitude of these differences, however, varies considerably between platforms and characteristics studied. Nonetheless, for most constructs and platforms we find low significant positive correlations between retrospective and aggregated in-situ values" (p. 143).

Verbeij et al. (2021) examined the accuracy and convergent validity of retrospective surveys and ESM surveys, by comparing adolescents' responses to these self-report measures with their digital trace data (criterion) (see Table 4.10). In both retrospective surveys and ESM, adolescents overestimated their time spent on social media. They more accurately estimated their time spent on platforms that are used in a less fragmented way (Instagram) than on platforms that are used in a more fragmented way (Snapchat). The study showed minimum acceptable convergent (between-person) validity of these measures. The within-person convergent validity (ESM estimates) was unacceptable. Both between- and within-person convergent validity decreased over time (due to a fatigue effect).

Table 4.10 Examination of the accuracy and convergent validity of retrospective surveys and experience sampling method (ESM) surveys

Data *N* = 125 with Android smartphones (time spent), adolescents Experience Sampling Method measurement: an app installed on respondents' smartphones was programmed to generate six notifications per day for a period of three weeks. Adolescents' ESM estimates of their social media use were obtained by three questions per ESM assessment, in which adolescents were asked to estimate the time spent using Instagram, WhatsApp, and Snapchat in the previous hour. Digital trace data: an application on respondents' phone tracked their app usage (i.e., type of app and duration of use) during the three-week ESM period. Every 5 min, this application retrieved the Android log data on adolescents' personal devices. Records of app use when adolescents' screen was turned off were excluded (roughly 2% of app usage estimates). Retrospective social media use: at the last ESM assessment of each week, participants received three additional questions about the time they had spent with Instagram, WhatsApp, and Snapchat in the previous week. The researchers measured typical weekly time spent on social media using direct estimates that assessed the frequency and duration of adolescents' social media use. The variable time spent on social media in a typical week was calculated by multiplying the number of days on which adolescents typically use a specific platform by the total number of minutes they used these platforms on these days.

Source: Verbeij et al. (2021, p. 4).

Karnowski et al. (2019) tried to identify factors that impact the differences between retrospective and momentary duration values, but could not find effects from the variables they studied: frequency of use, habit strength, context stability, involvement, and social desirability.

Think-aloud

Alhadreti and Mayhew's (2018) study into the evaluation of a library website, showed that concurrent think-aloud outperformed the retrospective method in terms of higher positive ratings from participants, being faster, producing better outcomes while not suffering from reactivity (see also Alshammari et al., 2015). Schaap (2004), however, found in his study on viewers' interpretation of TV news that the thought-listing technique yields more material than the think-aloud method which he explained by an individual's limited capacity to perform multiple tasks; the ongoing stream of sounds and images from television news interferes with the verbalization task in the think-aloud method.

4.4 Measures of responses during media exposure

Cognitive responses — attention, involvement, and engagement — and affective responses (valence, arousal and emotions), and experiences during media exposure are considered relevant concepts in studies of media use and effects (see Chapter 2). Given the many different definitions and operationalizations, we can only give a few examples of how these concepts are measured. Alternative scales for attention, involvement, engagement, and experience can be found, for example, in Bearden et al. (2011) and McQuarrie et al. (2011).

Attention

Attention can be measured with several methods including behavioral observation (e.g., facial expressions, eye gaze, reaction times), psychophysiological methods (e.g., blood pressure, galvanic skin response, brain wave activity), and self-reports (Chaffee & Schleuder, 1986). We return to eye tracking measures in Chapter 7 and to psychophysiological methods in Chapter 8. Here we discuss self-reports.

Attention is generally measured as a continuum by asking participants directly how much attention they have paid to a medium or a message. Romantan et al. (2008) applied a single-item measure: "How much attention do you pay to information about health or medical topics … [on TV / on the radio / in newspapers / in magazines / on the internet]. Would you say a lot, some, a little, or not at all?" (p. 84).

The self-report (message attention) measure developed by Laczniak et al. (1989), included five items: How much attention did you pay to … , How much did you notice … , How much attention did you concentrate on … , How involved were you with … , How much thought did you put into evaluating … (answer categories (1) none … (7) very much). This measure successfully passed reliability and manipulation checks.

It is shown that exposure and attention are separate dimensions (e.g., Chaffee & Schleuder, 1986; Drew & Weaver, 1990; Fikkers et al., 2015; Slater & Rasinski, 2005) and that combining these two adds to the explanation of media effects. An interesting issue is to what extent attention is a general trait, or medium, genre or issue specific. Chaffee and Schleuder (1986) concluded that in their study attention was a general trait to a genre (news): "There is some evidence of fluctuation in attention from one medium to another, one kind of news to another, and one time to another, but these dimensions of variation are overshadowed by the general trait that we might call attentiveness to news media" (p. 102). These conclusions correspond to Eveland et al.'s (2009) conclusion that attention captures a general predisposition to news, "very closely related to — and possibly inseparable from political interest" (p. 240).

Despite concerns that measures of attention may be confused by interest and concern about the subject, Slater et al. (2009) concluded that self-reported attention in their study "cannot be dismissed as a mere proxy for, or consequence of, prior concerns about the topic of the stories" (p. 131).

Involvement

Zaichkowsky (1985) developed a "personal involvement inventory" with items such as important — unimportant, irrelevant-relevant, matters to me — doesn't matter, and essential-nonessential. A revised version is published in McQuarrie & Munson (1987).

Moorman et al. (2012) developed a scale with eight different statements drawn from previous studies on involvement (e.g., Bryant & Comisky 1978; Moorman et al., 2007; Norris & Colman, 1993): (1) I was fascinated by [the media segment]; (2) My thoughts wandered off during [the media segment]; (3) I thought [the medium segment] was exciting; (4) I was distracted during [the media segment]; (5) I thought the [medium segment] was boring; (6) I paid attention to [the medium segment]; (7) I thought of other things during the [medium segment]; and (8) I thought the [medium segment] was interesting (five-point scale: 1 = not at all; 5 = very much). In their study, the items loaded on one dimension that formed an internally valid scale.

Engagement

Engagement is also conceptualized and operationalized in many different ways, including cognitive, affective and behavioral aspects (see also Chapter 2). The Advertising Research Foundation identified no less than 25 definitions (Advertising Research Foundation, 2006a). Engagement is measured by repeat viewing, time spent, liking, sharing, word-of-mouth for instance (Advertising Research Foundation, 2006b, 2007; Brodie et al., 2011; Calder et al., 2009; Hollebeek, 2011; Napoli, 2012; Webster et al., 2014). We refer to Hollebeek et al. (2014), Obilo et al. (2020), and Araujo et al. (2022) for extensive considerations and suggestions for measuring engagement based on the investment of cognitive, emotional, behavioral, and social resources.

Valence, arousal, and emotional responses

Valence and arousal are usually measured with neurobiological measures (see Chapter 8), but self-reports are applied as well. For example, Osmundsen et al. (2022) measured "valence by asking on nine-point scales whether participants responded with 'Happy, positive feelings' or 'Unhappy, negative feelings' when viewing that image." They "measured arousal by asking, on nine-point scales, whether participants had 'no reaction' or a 'strong reaction' when viewing the image" (p. 57). Yzer et al. (2011) developed self-report measures for these two concepts which are summarized in Table 4.11.

Table 4.11 Real time valence and arousal ratings using self-reports

"Participants used a computer mouse to rate each ad on a moment-by-moment basis. A horizontal line at the bottom of the screen gave participants visual feedback of the cursor's position. The computer logged the average of 10 measurements per second to indicate the cursor position at each second of the particular ad. The positions reflected a 7-point valence or arousal scale, ranging from 0 to 6.

We explained the valence and arousal dimensions using affective states that exemplify the two dimensions (e.g., Russell 1980, 2003). Participants in the valence condition were told about feeling happy versus unhappy to illustrate the task: 'You will be asked to describe your feelings along the dimension: happy vs. unhappy. At the right end of the scale you are happy, pleased, satisfied, contented, hopeful. At the left end of the scale is the opposite feeling.' The momentary rating task for participants in the arousal condition described the arousal dimension as: 'You will be asked to describe your feelings along the dimension: "stirred up" vs. bored. At the right end of the scale you are stirred up, stimulated, excited, frenzied, jittery, wide-awake, aroused. At the left end of the scale is the opposite feeling.'"

Source: Yzer et al. (2011, pp. 281–282).

Several measures for self-reported emotional responses are listed by Nabi (2007, 2009), including the well-known Positive and Negative Affect Schedule (PANAS) that consists of two 10-items mood scales (Watson et al., 1988). The scales include items such as "upset," "scared," "inspired," and "active."

Experiences

Malthouse et al. (2007) measured (magazine) experiences with 39 factors, each consisting of multiple items. Some examples of these factors are "The stories absorb me," "I often reflect on it," "I'm inspired," "I'm touched," "I build relationships by talking about and sharing it." For example, the "I'm touched factor" consisted of four items: "It helps me to see that there are good people in the world," "Some articles touch me deep down," "It features people who make you proud," and "The magazine definitely affects me emotionally."

Calder et al. (2016) distinguished five experience categories: interaction (to connect with others), discovery (to gain insights), transportation (to escape), identity (to express identity), and civic orientation (to contribute to society).

Voorveld et al. (2018) based their measurement on Bronner & Neijens (2006) with experience dimensions such as "information" (e.g., gave me useful information), "entertainment" (e.g., gave me enjoyment), "identification" (e.g., I recognized myself in it), and "social interaction" (e.g., enabled me to do or share something with others). Central to their approach was the absence of forced exposure, they concentrated on the engagement experience of users at a specific, recent, media consumption moment. Their study (n = 1,346, aged 13 and older) included Facebook, YouTube, LinkedIn, Twitter, Google, Instagram, Pinterest, and Snapchat, and showed how social media experiences are highly context specific: each social medium platform was experienced in a unique way.

4.5 Concluding remarks

Self-reports measures of media exposure are popular, have great advantages, but research also shows that self-report measures have serious conceptual, response, validity, and reliability issues. The studies that have been performed provide insight into how the characteristics of the self-reports, such as type of recall, unit of measure, reference period, response scale, help, and linkage options, influence their performance. Given the many factors involved, and the heterogeneity of the methodological studies (which differ in type of medium, content, platform, population and media landscape), definitive, general conclusions about the best way to ask media users about their media exposure are not possible. However, the studies provide invaluable insights useful for evaluating the quality of self-report measures in specific studies, and for making trade-offs when deciding which measure to apply. Making this process of selecting exposure measures transparent and explicitly evaluating pros and cons of different measures is recommended. In addition, further research on the quality of self-reports is needed, preferably embedded in systematic research programs. We will come back to this in Chapter 10.

References

Advertising Research Foundation (2006a). *Engagement: Definitions and anatomy*. Advertising Research Foundation.

Advertising Research Foundation (2006b). *Measures of engagement*. Advertising Research Foundation.

Advertising Research Foundation (2007). *Measures of engagement: Volume II*. Advertising Research Foundation.

Alhadreti, O., & Mayhew, P. (2018). Rethinking thinking aloud: A comparison of three Think-Aloud Protocols. *Proceedings of the 2018 CHI Conference on Human Factors in Computing Systems, USA*, Article 44. https://doi.org/10.1145/3173574.3173618

Alshammari, T., Alhadreti, O., & Mayhew, P. (2015). When to ask participants to think aloud: A comparative study of concurrent and retrospective think-aloud methods. *International Journal of Human Computer Interaction*, *6*(3), 48–64. https://www.cscjournals.org/library/manuscriptinfo.php?mc=IJHCI-118

Althaus, S. L., & Tewksbury, D. H. (2007). *Toward a new generation of media use measures for the ANES* (Report No. nes011903). American National Election Studies. https://electionstudies.org/wp-content/uploads/2018/04/nes011903.pdf

Andersen, K., de Vreese, C. H., & Albæk, E. (2016). Measuring media diet in a high-choice environment—Testing the List-Frequency Technique. *Communication Methods and Measures*, *10*(2–3), 81–98. https://doi.org/10.1080/19312458.2016.1150973

Anderson, C. A., & Dill, K. E. (2000). Video games and aggressive thoughts, feelings, and behavior in the laboratory and in life. *Journal of Personality and Social Psychology*, *78*(4), 772–790. https://doi.org/10.1037//0022-3514.78.4.772

Araujo, T., Ausloos, J., van Atteveldt, W., Loecherbach, F., Moeller, J., Ohme, J., Trilling, D., van de Velde, B., de Vreese, C., & Welbers, K. (2022). OSD2F: An open-source data donation framework. *Computational Communication Research*, *4*(2), 372–387. https://doi.org/10.5117/CCR2022.2.001.ARAU

Araujo, T., Wonneberger, A., Neijens, P., & de Vreese, C. (2017). How much time do you spend online? Understanding and improving the accuracy of self-reported measures of internet use. *Communication Methods & Measures*, *11*(3), 173–190. https://doi.org/10.1080/19312458.2017.1317337

Avery, R. J., Eisenberg, M. D., & Simon, K. I. (2012). The impact of direct-to-consumer television and magazine advertising on antidepressant use. *Journal of Health Economics*, *31*(5), 705–718. https://doi.org/10.1016/j.jhealeco.2012.05.002

Bachl, M., & Scharkow, M. (2019). Some suggestions on dealing with measurement error in linkage analyses. In C. Peter, T. Naab, & R. Kühne (Eds.), *Methoden und forschungslogik der kommunikationswissenschaft: Vol. 14. Measuring media use* *and exposure: Recent developments and challenges* (pp. 175–195). Herbert von Halem Verlag.

Bearden, W. O., Netemeyer, R. G., & Haws, K. L. (2011). *Handbook of marketing scales: Multi-item measures for marketing and consumer behavior research*. Sage. https://doi.org/10.4135/9781412996761

Bechtel, R. B., Achelpohl, C., & Akers, R. (1972). Correlates between observed behavior and questionnaire responses on television viewing. In E. A. Rubinstein, G. A. Comstock, & J. P. Murray (Eds.), *Television and human behavior: A technical report to the surgeon general's advisory committee on television and social behavior: Vol. 4* (pp. 274–344). Government Printing Office.

Belli, R. (1998). The structure of autobiographical memory and the event history calendar: Potential improvements in the quality of retrospective reports in surveys. *Memory*, *6*(4), 383–406. https://doi.org/10.1080/741942610

Belli, R. F., Schwarz, N., Singer, E., & Talarico, J. (2000). Decomposition can harm accuracy of behavioral frequency reports. *Applied Cognitive Psychology*, *14*(4), 295–308. https://doi.org/10.1002/1099-0720(200007/08)14:4<295::AID-ACP646>3.0.CO;2-1

Belson, W. A. (1981). *The design and understanding of survey questions*. Gower.

Belson, W. A. (1986). *Validity in survey research*. Gower.

Benbunan-Fich, R. (2001). Using protocol analysis to evaluate the usability of a commercial web site. *Information Management*, *39*(2), 151–163. https://doi.org/10.1016/S0378-7206(01)00085-4

Bergkvist, L., & Langner, T. (2020). Perspectives: Four steps toward more valid and comparable self-report measures in advertising research. *International Journal of Advertising*, *39*(5), 738–755. https://doi.org/10.1080/02650487.2019.1665398

Bij de Vaate, N. A. J. D., Valkenburg, P. M., Janssen, L. H. C., van der Wal, A., Keijsers, L., & Beyens, I. (2023). *Social media self-effects on adolescents' self-esteem: A 100-day diary study*. Psyarxiv. https://doi.org/10.31234/osf.io/qre85

Biocca, F., David, P., & West, M. (1994). Continuous response measurement (CRM): A computerized tool for research on the cognitive processing of communication messages. In A. Lang (Ed.), *Communication series: Measuring psychological responses to media messages* (pp. 15–64). Erlbaum.

Blair, E., & Burton, S. (1987). Cognitive processes used by survey respondents to answer behavioral frequency questions. *Journal of Consumer Research*, *14*(2), 280–288. https://doi.org/10.1086/209112

Boase, J., & Ling, R. S. (2013). Measuring mobile phone use: Self-report versus log data. *Journal of Computer-Mediated Communication*, *18*(4), 508–519. https://doi.org/10.1111/jcc4.12021

Bradburn, N. M., & Sudman, S. (1979). *Improving interview method and questionnaire design*. Jossey-Bass.

Bradburn, N. M., Rips, L. J., & Shevell, S. K. (1987). Answering autobiographical questions: The impact of memory and inference on surveys. *Science*, *236*(4798), 157–161. https://doi.org/10.1126/science.3563494

Branch, J. L. (2000). Investigating the information-seeking processes of adolescents: The value of using think alouds and think afters. *Library Information Science Research*, *22*(4), 371–392. https://doi.org/10.1016/S0740-8188(00)00051-7

Brodie, R., Hollebeek, L. D., Ilic, A., & Juric, B. (2011). Customer engagement: Conceptual domain, fundamental propositions & implications for research in service marketing. *Journal of Service Research*, *14*(3), 252–271. https://doi.org/10.1177/1094670511411703

Bronner, F., & Neijens, P. (2006). Audience experiences of media context and embedded advertising: A comparison of eight media. *International Journal of Market Research*, *48*(1), 81–100. https://doi.org/10.1177/147078530604800106

Brown, M. (1990). *Dear reader: Some readership measurement questions … and some answers*. Research Services Ltd & the British Market Research Bureau Ltd.

Bryant, J., & Comisky, P. W. (1978). The effect of positioning a message within differentially cognitively involving portions of a television segment on recall of the message. *Human Communication Research*, *5*(1), 63–75. https://doi.org/10.1111/j.1468-2958.1978.tb00623.x

Bülow, A., Janssen, L., Dietvorst, E., Bij de Vaate, N., Hillegers, M., Valkenburg, P., & Keijsers, L. (2024). *From burden to enjoyment: A user-centered approach to engage adolescents in intensive longitudinal research*. https://doi.org/10.31234/osf.io/rnuvz.

Burnell, K., George, M. J., Kurup, A. R., Underwood, M. K., & Ackerman, R. A. (2021). Associations between self-reports and device-reports of social networking site use: An application of the truth and bias model. *Communication Methods and Measures*, *15*(2), 156–163. https://doi.org/10.1080/19312458.2021.1918654

Burton, S., & Blair, E. (1991). Task conditions, response formulation processes, and response accuracy for behavioral frequency questions in surveys. *Public Opinion Quarterly*, *55*(1), 50–79. https://doi.org/10.1086/269241

Calder, B. J., Isaac, M. S., & Malthouse, E. C. (2016). How to capture consumer experiences: A context-specific approach to measuring engagement predicting consumer behavior across qualitatively different experiences. *Journal of Advertising Research*, *56*(1), 39–52. https://doi.org/10.2501/JAR-2015-028

Calder, B. J., Malthouse, E. C., & Schaedel, U. (2009). An experimental study of the relationship between online engagement and advertising effectiveness. *Journal of Interactive Marketing*, *23*(4), 321–331. https://doi.org/10.1016/j.intmar.2009.07.002

Cannell, C., Miller, P., & Oksenberg, L. (1981). Research on interviewing techniques. In S. Leinhardt (Ed.), *Sociological methodology 1981* (pp. 389–437). Jossey-Bass.

Cernat, A., Keusch, F., Bach, R. L., & Pankowska, P. K. (2024). Estimating measurement quality in digital trace data and surveys using the MultiTrait MultiMethod Model. *Social Science Computer Review*, 0(0). https://doi.org/10.1177/08944393241254464

Chaffee, S. H., & Schleuder, J. (1986). Measurement and effects of attention to media news. *Human Communication Research*, *13*(1), 76–107. https://doi.org/10.1111/j.1468-2958.1986.tb00096.x

Chang, L., & Krosnick, J. A. (2003). Measuring the frequency of regular behaviors: Comparing the 'typical week' to the 'past week'. *Sociological Methodology*, *33*(1), 55–80. https://doi.org/10.1111/j.0081-1750.2003.t01-1-00127.x

Clancy, K. J., Ostlund, L. E., & Wyner, G. A. (1979). False reporting of magazine readership. *Journal of Advertising Research*, *19*(5), 23–30.

Clay, R., Barber, J. M., & Shook, N. J. (2013). Techniques for measuring selective exposure: A critical review. *Communication Methods and Measures*, *7*(3–4), 147–171. https://doi.org/10.1080/19312458.2013.813925

Cohen, A. A., & Lemish, D. (2003). Real time and recall measures of cell phone use: Some methodological concerns and empirical applications. *New Media & Society*, *5*(2), 167–183. https://doi.org/10.1177/1461444803005002002

Converse, J. M., & Presser, S. (1986). *Survey questions: Hand crafting the standardized questionnaire*. Sage.

Corlett, T. (1964). The IPA National Readership Survey: Some problems and possible solutions. *Journal of Advertising Research*, *4*(4), 4–10.

Corlett, T., & Osborne, D. W. (1966). *The development of reading frequency scales*. Institute of Practitioners in Advertising.

Coromina, L., & Saris, W. E. (2009). Quality of media use measurement. *International Journal of Public Opinion Research*, *21*(4), 424–450. https://doi.org/10.1093/ijpor/edp014

Csikszentmihalyi, M., & Larson, R. (1987). Validity and reliability of the experience-sampling method. *The Journal of Nervous & Mental Disease*, *175*(9), 526–536. https://doi.org/10.1097/00005053-198709000-00004

de Vreese, C. H., Boukes, M., Schuck, A., Vliegenthart, R., Bos, L., & Lelkes, Y. (2017). Linking survey and media content data: Opportunities, considerations, and pitfalls. *Communication Methods and Measures*, *11*(4), 221–244. https://doi.org/10.1080/19312458.2017.1380175

DeCastellarnau, A. (2018). A classification of response scale characteristics that affect data quality: A literature review. *Quality and Quantity*, *52*, 1523–1559. https://doi.org/10.1007/s11135-017-0533-4

Deng, T., Kanthawala, S., Meng, J., Peng, W., Kononova, A., Hao, Q., Zhang, Q., & Prabu, D. (2019). Measuring smartphone usage and task switching with log tracking and self-reports. *Mobile Media & Communication*, *7*(1), 3–23. https://doi.org/10.1177/2050157918761491

Dilliplane, S., Goldman, S. K., & Mutz, D. C. (2013). Televised exposure to politics: New measures for a fragmented media environment. *American Journal of Political Science*, *57*(1), 236–248. https://doi.org/10.1111/j.1540-5907.2012.00600.x

Drew, D., & Weaver, D. (1990). Media attention, media exposure, and media effects. *Journalism Quarterly*, *67*(4),740–748. https://doi.org:10.1177/107769909006700428

Ericsson, K. A., & Simon, H. A. (1993). *Protocol analysis: Verbal reports as data.* MIT Press.

Ernala, S. K., Burke, M., Leavitt, A., & Ellison, N. B. (2020). How well do people report time spent on Facebook? An evaluation of established survey questions with recommendations. *Proceedings of the 2020 CHI Conference on Human Factors in Computing Systems, USA*, 1–14. https://doi.org/10.1145/3313831.3376435

Eveland, W. P., & Dunwoody, S. (2000). Examining information processing on the World Wide Web using think aloud protocols. *Media Psychology*, *2*(3), 219–244. https://doi.org/10.1207/S1532785XMEP0203_2

Eveland, W., Hutchens, M., & Shen, F. (2009). Exposure, attention, or "use" of news? Assessing aspects of the reliability and validity of a central concept in political communication research. *Communication Methods & Measures*, *3*(4), 223–244. https://doi.org/10.1080/19312450903378925

Fikkers, K. M., Piotrowski, J. T., & Valkenburg, P. M. (2015). Assessing the reliability and validity of television and game violence exposure measures. *Communication Research*, *44*(1), 117–143. https://doi.org/10.1177/0093650215573863

Fisher, C. D., & To, M. L. (2012). Using experience sampling methodology in organizational behavior. *Journal of Organizational Behavior, 33*(7), 865–877. https://doi.org/10.1002/job.1803

Fortin, C., & Rousseau, R. (1998). Interference from short-term memory processing on encoding and reproducing brief durations. *Psychological Research*, *61*(4), 269–276. https://doi.org/10.1007/s004260050031

Gabriel, A. S., Podsakoff, N. P., Beal, D. J., Scott, B. A., Sonnentag, S., Trougakos, J. P., & Butts, M. M. (2019). Experience sampling methods: A discussion of critical trends and considerations for scholarly advancement. *Organizational Research Methods*, *22*(4), 969–1006. https://doi.org/10.1177/1094428118802626

Galinat, W. H., & Borg, I. (1987). On symbolic temporal information: Beliefs about the experience of duration. *Memory and Cognition*, *15*(4), 308–317. https://doi.org/10.3758/BF03197033

Greenberg, B. S., Dervin, B., & Dominick, J. (1968). Do people watch "television" or "programs"? A measurement problem. *Journal of Broadcasting*, *12*(4), 367–376. https://doi.org/10.1080/08838156809386266

Griffioen, N., van Rooij, M. M. J. W., Lichtwarck-Aschoff, A., & Granic, I. (2020a). A stimulated recall method for the improved assessment of quantity and quality of social media use. *Journal of Medical Internet Research*, *22*(1), Article e15529. https://doi.org/10.2196/15529

Griffioen, N., van Rooij, M. M. J. W., Lichtwarck-Aschoff, A., & Granic, I. (2020b). Toward improved methods in social media research. *Technology, Mind, and Behavior*, *1*(1). https://doi.org/10.1037/tmb0000005

Guess, A. (2015). Measure for measure: An experimental test of online political media exposure. *Political Analysis*, *23*(1), 59–75. https://doi.org/10.1093/pan/mpu010

Haenschen, K. (2020). Self-reported versus digitally recorded: Measuring political activity on Facebook. *Social Science Computer Review*, *38*(5), 567–583. https://doi.org/10.1177/0894439318813586

Hobson, J. W. (1956). *The selection of advertising media* (2nd ed.). Business Publications on behalf of the Institute of Practitioners in Advertising.

Hollebeek, L. (2011). Exploring customer brand engagement: Definition and themes. *Journal of Strategic Marketing*, *19*(7), 555–573. https://doi.org/10.1080/0965254X.2011.599493

Hollebeek, L. D., Glynn, M. S., & Brodie, R. J. (2014). Consumer brand engagement in social media: Conceptualization, scale development and validation. *Journal of Interactive Marketing*, *28*(2), 149–65. https://doi.org/10.1016/j.intmar.2013.12.002

Hoppmann, T. K. (2009). Examining the 'point of frustration': The think-aloud method applied to online search tasks. *Quality & Quantity*, *43*, 211–224. https://doi.org/10.1007/s11135-007-9116-0

Hornik, J. (1984). Subjective vs. objective time measures: A note on the perception of time in consumer behavior. *Journal of Consumer Research*, *11*(1), 615–618. https://doi.org/10.1086/208998

Hufford, M. R., Shiffman, S., Paty, J., & Stone, A. A. (2001). Ecological momentary assessment: Real-world, real-time measurement of patient experience. In J. Fahrenberg & M. Myrtek (Eds.), *Progress in ambulatory assessment: Computer-assisted psychological and psychophysiological methods in monitoring and field studies* (pp. 69–92). Hogrefe & Huber.

Jerit, J., Barabas, J., Pollock, W., Banducci, S., Stevens, D., & Schoonvelde, M. (2016). Manipulated vs. measured: Using an experimental benchmark to investigate the performance of self-reported media exposure. *Communication Methods and Measures*, *10*(2–3), 99–114. https://doi.org/10.1080/19312458.2016.1150444

Johannes, N., Nguyen, T.-V., Weinstein, N., & Przybylski, A. K. (2021). Objective, subjective, and accurate reporting of social media use: No evidence that daily social media use correlates with personality traits, motivational states, or well-being. *Technology, Mind, and Behavior*, *2*(2). https://doi.org/10.1037/tmb0000035

Junco, R. (2013). Comparing actual and self-reported measures of Facebook use. *Computers in Human Behavior*, *29*(3), 626–631. https://doi.org/10.1016/j.chb.2012.11.007

Jurgens, P., Stark, B., & Magin, M. (2019). Two half-truths make a whole? On bias in self-reports and tracking data. *Social Science Computer Review*, *38*(5), 600–615. https://doi.org/10.1177/0894439319831643

Juster, F. T., Ono, H., & Stafford, F. P. (2003). An assessment of alternative measures of time use. *Sociological Methodology, 33*, 19–54. https://www.jstor.org/stable/1519852

Juster, F. T., & Stafford, F. P. (1991). The allocation of time: Empirical findings, behavioral models, and problems of measurement. *Journal of Economic Literature, 29*(2), 471–522. https://www.jstor.org/stable/2727521

Kahneman D., Krueger A. B, Schkade D. A., Schwarz, N, & Stone A. A. (2004). A survey method for characterizing daily life experience: The day reconstruction method. *Science, 306*(5702), 1776–1780. https://doi.org/10.1126/science.1103572

Karnowski, V., Naab. T. K., & Schlutz (2019). On the challenges of measuring mobile social media use: Explaining differences between data from surveys and mobile experience sampling. In C. Peter, T. Naab, & R. Kühne (Eds.), *Methoden und forschungslogik der kommunikationswissenschaft: Vol. 14. Measuring media use and exposure: Recent developments and challenges* (pp. 114–138). Herbert von Halem Verlag.

Kaye, L. K., Orben, A., Ellis, D. A., Hunter, S. C., & Houghton, S. (2020). The conceptual and methodological mayhem of "Screen Time." *International Journal of Environmental Research and Public Health, 17*(10), Article 3661. https://doi.org/10.3390/ijerph17103661

Kobayashi, T., & Boase, J. (2012). No such effect? The implications of measurement error in self-report measures of mobile communication use. *Communication Methods and Measures, 6*(2), 126–143. https://doi.org/10.1080/19312458.2012.679243

Krosnick, J. A. (1991). Response strategies for coping with the cognitive demands of attitude measures in surveys. *Applied Cognitive Psychology, 5*(3), 213–236. https://doi.org/10.1002/acp.2350050305

Kubey, R., Larson, R., & Csikszentmihalyi, M. (1996). Experience sampling method applications to communication research questions. *Journal of Communication, 46*(2), 99–120. https://doi.org/10.1111/j.1460-2466.1996.tb01476.x

Laczniak, R. N., Muehling, D. D., & Grossbart, S. (1989). Manipulating message involvement in advertising research. *Journal of Advertising, 18*(2), 28–38. https://doi.org/10.1080/00913367.1989.10673149

Larson, R., & Csikszentmihalyi, M. (1983). The experience sampling method. *New Directions for Methodology of Social and Behavioral Science, 15*, 41–56.

Lee, C., Hornik, R., & Hennessy, M. (2008). The reliability and stability of general media exposure measures. *Communication Methods and Measures, 2*(1–2), 6–22. http://dx.doi.org/10.1080/19312450802063024

Lin, Y. H., Lin, Y. C., Lee, Y. H., Lin, P. H., Lin, S. H., Chang, L. R., Tseng, H. W., Yen, L. Y., Yang, C. C. H., & Kuo, T. B. J. (2015). Time distortion associated with smartphone addiction: Identifying smartphone addiction via a mobile application (App). *Journal of Psychiatric Research, 65*, 139–145. https://doi.org/10.1016/j.jpsychires.2015.04.003

Link, M. W., Murphy, J., Schober, M. F., Buskirk, T. D., Childs, J. H., & Tesfaye, C. L. (2014). Mobile technologies for conducting, augmenting and potentially replacing surveys: Report of the AAPOR task force on emerging technologies in public opinion research. *Public Opinion Quarterly*, *78*(4), 779–787. https://doi.org/10.1093/poq/nfu054

Malthouse, E. C., Calder, B. J., & Tamhane, A. (2007). The effects of media context experiences on advertising effectiveness. *Journal of Advertising*, *36*(3), 7–18. https://doi.org/10.2753/JOA0091-3367360301

Marder, E. (1967). How good is the editorial-interest method of measuring magazine audiences?. *Journal of Advertising Research*, *7*(1), 2–6. http://www.jstor.org/stable/2351259

McGlathery, D. (1967). Claimed frequency vs. editorial-interest measures of repeat magazine audiences. *Journal of Advertising Research*, *7*(1), 7–15.

McQuarrie, E. F., & Munson, J. M. (1987). The Zaichkowsky personal involvement inventory: Modification and extension. In M. Wallendorf & P. Anderson (Eds.), *NA Advances in consumer research, Vol. 14* (pp. 36–40). Association for Consumer Research.

McQuarrie, W. O., Netemeyer, R. G., & Haws, K. L. (2011). *Handbook of marketing scales: Multi-item measures for marketing and consumer behavior research*. Sage.

Menon, G., & Yorkston, E. A. (2000). The use of memory and contextual cues in the formation of behavioral frequency judgments. In A. Stone, J. Turkkan, C. Bachrach, J. Jobe, H. Kurtzman, & V. S. Cain (Eds.), *The science of self-report: Implications for research and practice* (pp. 63–79). Erlbaum.

Moorman, M., Neijens, P. C., & Smit, E. G. (2007). The effects of program involvement on commercial exposure and recall in a naturalistic setting. *Journal of Advertising*, *36*(1), 125–141. https://doi.org/10.2753/JOA0091-3367360109

Moorman, M., Willemsen, L., Neijens, P. C., & Smit, E. G. (2012). Program-involvement effects on commercial attention and recall of successive and embedded advertising. *Journal of Advertising*, *41*(2), 25–37. https://doi.org/10.2753/JOA0091-3367410202

Mutz, D., & Young, L. (2011). Communication and public opinion. *Public Opinion Quarterly*, *75*(5), 1018–1044. https://doi.org/10.1093/poq/nfr052

Naab, T. K., Karnowski, V., & Schlütz, D. (2019). Reporting mobile social media use: How survey and experience sampling measures differ. *Communication Methods and Measures*, *13*(2), 126–147. https://doi.org/10.1080/19312458.2018.1555799

Nabi, R. L. (2007). Emotion and persuasion: A social cognitive perspective. In D. R. Roskos-Ewoldsen & J. Monahan (Eds.), *Social cognition and communication: Theories and methods* (pp. 377–398). Erlbaum.

Nabi, R. L. (2009). Emotion and media effects. In R. L. Nabi & M. B. Oliver (Eds.), *The Sage handbook of media processes and effects* (pp. 205–221). Sage.

Napoli, P. M. (2012). Audience evolution and the future of audience research. *International Journal on Media Management*, *14*(2), 79–97. https://doi.org/10.1080/14241277.2012.675753

Neter, J., & Waksberg, J. (1964). A study of response errors in expenditures data from household interviews. *Journal of the American Statistical Association*, *59*(305), 18–55. https://doi.org/10.1080/01621459.1964.10480699

Niederdeppe, J. (2014). Conceptual, empirical, and practical issues in developing valid measures of public communication campaign exposure. *Communication Methods and Measures*, *8*(2), 138–161. https://doi.org/10.1080/19312458.2014.903391

Niederdeppe, J. (2016). Meeting the challenge of measuring communication exposure in the digital age. *Communication Methods and Measures*, *10*(2–3), 170–172. https://doi.org/10.1080/19312458.2016.1150970

Nielsen, J. (1994). *Usability engineering*. Elsevier.

Norris, C. E., & Colman, A. M. (1993). Context effects on memory for television advertisements. *Social Behavior and Personality*, *21*(4), 279–286. https://doi.org/10.2224/sbp.1993.21.4.279

Obilo, O. O., Chefor, E., & Saleh, A. (2020). Revisiting the consumer brand engagement concept. *Journal of Business Research*, *126*, 634–643. https://doi.org/10.1016/j.jbusres.2019.12.023

Ohme, J., Araujo, T., de Vreese, C. H., & Piotrowski, J. T. (2021). Mobile data donations: Assessing self-report accuracy and sample biases with the iOS Screen Time function. *Mobile Media & Communication*, *9*(2), 293–313. https://doi.org/10.1177/2050157920959106

Opree, S., Buijzen, M., & van Reijmersdal, E. (2021). Reliability and validity of children's advertising exposure measures. *Young Consumers*, *22*(4), 660–674. https://doi.org/10.1108/YC-10-2020-1243

Osmundsen, M., Hendry, D. J., Laustsen, L., Smith, K. B., & Petersen, M. B. (2022). The psychophysiology of political ideology: Replications, reanalyses, and recommendations. *The Journal of Politics*, *84*(1), 50–66. https://doi.org/10.1086/714780

Otto, L. P., Loecherbach, F., & Vliegenthart, R. (2023). Linkage analysis revised—Linking digital traces and survey data. *Communication Methods and Measures*. Advance online publication. https://doi.org/10.1080/19312458.2023.2257595

Otto, L. P., Thomas, F., Glogger, I., & de Vreese, C. H. (2022). Linking media content and survey data in a dynamic and digital media environment—mobile longitudinal linkage analysis. *Digital Journalism*, *10*(1), 200–215. https://doi.org/10.1080/21670811.2021.1890169

Parry, D., Davidson, B. I., Sewall, C., Fisher, J. T., Mieczkowski, H., & Quintana, D. S. (2021). A systematic review and meta-analysis of discrepancies between logged and self-reported digital media use. *Nature Human Behavior*, *5*(11), 1535–1547. https://doi.org/10.1038/s41562-021-01117-5

Parslow, R. C., Hepworth, S. J., & McKinney, P. A. (2003). Recall of past use of mobile phone handsets. *Radiation Protection Dosimetry*, *106*(3), 233–240. https://doi.org/10.1093/oxfordjournals.rpd.a006354

Potts, R., & Seger, J. (2013). Validity of adults' retrospective memory for early television viewing. *Communication Methods and Measures*, *7*(1), 1–25. https://doi.org/10.1080/19312458.2012.760731

Price, V. (1993). The impact of varying reference periods in survey questions about media use. *Journalism Quarterly*, *70*(3), 615–27. https://doi.org/10.1177/107769909307000312

Prior, M. (2009a). Improving media effects research through better measurement of news exposure. *Journal of Politics*, *71*(3), 893–908. https://doi.org/10.1017/S0022381609090781

Prior, M. (2009b). The immensely inflated news audience: Assessing bias in self-reported news exposure. *Public Opinion Quarterly*, *73*(1), 130–143. https://doi.org/10.1093/poq/nfp002

Prior, M. (2013). The challenge of measuring media exposure: Reply to Dilliplane, Goldman, and Mutz. *Political Communication*, *30*(4), 620–634. https://doi.org/10.1080/10584609.2013.819539

Rintala, A., Wampers, M., Myin-Germeys, I., & Viechtbauer, W. (2019). Response compliance and predictors thereof in studies using the experience sampling method. *Psychological Assessment*, *31*(2), 226–235. https://doi.org/10.1037/pas0000662

Rintala, A., Wampers, M., Myin-Germeys, I., & Viechtbauer, W. (2020). Momentary predictors of compliance in studies using the experience sampling method. *Psychiatry Research*, *286*, Article 112896. https://doi.org/10.1016/j.psychres.2020.112896

Robinson, M. D., & Clore, G. L. (2002). Belief and feeling: Evidence for an accessibility model of emotional self-report. *Psychological Bulletin*, *128*(6), 934–960. https://doi.org/10.1037/0033-2909.128.6.934

Romantan, A., Hornik, R., Price, V., Cappella, J., & Viswanath, K. (2008). A comparative analysis of the performance of alternative measures of exposure. *Communication Methods and Measures*, *2*(1–2), 80–99. https://doi.org/10.1080/19312450802062539

Rozendaal, E., Buijzen, M., & Valkenburg, P. M. (2012). Think-Aloud process superior to thought-listing in increasing children's critical processing of advertising. *Human Communication Research*, *38*(2), 199–221. https://doi.org/10.1111/j.1468-2958.2011.01425.x

Russell, J. A. (1980). A circumplex model of affect. *Journal of Personality and Social Psychology*, *39*(6), 1161–1178. https://doi.org/10.1037/h0077714

Russell, J. A. (2003). Core affect and the psychological construction of emotion. *Psychological Review*, *110*(1), 145–172. https://doi.org/10.1037/0033-295X.110.1.145

Schaap, G. (2004). Using protocol analysis in television news research: Proposal and first tests. In K. Renckstorf, D. McQuail, J. Rosenbaum, & G. Schaap (Eds.), *Action theory and communication research* (pp. 115–140). Mouton de Gruyter.
Scharkow, M. (2016). The accuracy of self-reported internet use—A validation study using client log data. *Communication Methods and Measures*, *10*(1), 13–27. https://doi.org/10.1080/19312458.2015.1118446
Scharkow, M. (2019). The reliability and temporal stability of self-reported media exposure: A meta-analysis. *Communication Methods and Measures*, *13*(3), 198–211. https://doi.org/10.1080/19312458.2019.1594742
Schnauber-Stockmann, A., & Karnowski, V. (2020). Mobile devices as tools for media and communication research: A scoping review on collecting self-report data in repeated measurement designs. *Communication Methods and Measures*, *14*(3), 145–164. https://doi.org/10.1080/19312458.2020.1784402
Schuck, A. R. T., & de Vreese, C. H. (2008). The Dutch no to the EU constitution: Assessing the role of EU skepticism and the campaign. *Journal of Elections, Public Opinion and Parties*, *18*(1), 101–128. https://doi.org/10.1080/17457280701858656
Schwarz, N. (1999). Self-Reports: How the questions shape the answers. *American Psychologist*, *54*(2), 93–105. https://doi.org/10.1037/0003-066X.54.2.93
Schwarz, N. (2007). Retrospective and concurrent self-reports: The rationale for real-time data capture. In A. Stone, S. S. Shiffman, A. Atienza, & L. Nebeling (Eds.), *The science of real-time data capture: Self-reports in health research* (pp. 11–26). Oxford University Press.
Schwarz, N., Hippler, H.-J., Deutsch, B., & Stark, F. (1985). Response scales: Effects of category range on reported behavior and comparative judgments. *Public Opinion Quarterly*, *49*(3), 388–395. https://doi.org/10.1086/268936
Schwarz, N., & Oyserman, D. (2001). Asking questions about behavior: Cognition, communication and questionnaire construction. *American Journal of Evaluation*, *22*(2), 127–160. https://doi.org/10.1177/109821400102200202
Scollon, C. N., Kim-Prieto, C., & Diener, E. (2003). Experience sampling: Promises and pitfalls, strengths and weaknesses. *Journal of Happiness Studies*, *4*, 5–34. https://doi.org/10.1023/A:1023605205115
Sekely, W. S., & Blakney, V. L. (1994). The effect of response position on trade magazine readership and usage. *Journal of Advertising Research*, *34*(6), 53–59.
Sewall, C. J. R., Bear, T. M., Merranko, J., & Rosen, D. (2020). How psychosocial well-being and usage amount predict inaccuracies in retrospective estimates of digital technology use. *Mobile Media Communication*, *8*(3), 379–399. https://doi.org/10.1177/2050157920902830
Shapiro, M. A. (1994a). Signal detection measures of recognition memory. In A. Lang (Ed.), *Measuring psychological responses to the media* (pp. 133–148). Erlbaum.

Shapiro, M. A. (1994b). Think-aloud and thought-list procedures in investigating mental processes. In A. Lang (Ed.), *Measuring psychological responses to media messages* (pp. 1–14). Erlbaum.

Shiffman, S., Stone, A. A., & Hufford, M. R. (2008). Ecological momentary assessment. *Annual Review of Clinical Psychology, 4*(1), 1–32. https://doi.org/10.1146/annurev.clinpsy.3.022806.091415

Slater, M. D. (2004). Operationalizing and analyzing exposure: The foundation of media effects research. *Journalism & Mass Communication Quarterly, 81*(1), 168–183. https://doi.org/10.1177/107769900408100112

Slater, M. D. (2014). Reinforcing spirals model: Conceptualizing the relationship between media content exposure and the development and maintenance of attitudes. *Media Psychology, 18*(3), 370–395. https://doi.org/10.1080/15213269.2014.897236

Slater, M. D., Goodall, C. E., & Hayes, A. F. (2009). Self-reported news attention does assess differential processing of media content: An experiment on risk perceptions utilizing a random sample of U.S. local crime and accident news. *Journal of Communication, 59*(1), 117–134. https://doi.org/10.1111/j.1460-2466.2008.01407.x

Slater, M., & Rasinski, K. A. (2005). Media exposure and attention as mediating variables influencing social risk judgments. *Journal of Communication, 55*(4), 810–827. https://doi.org/10.1111/j.1460-2466.2005.tb03024.x

Smit, E. G., & Neijens, P. C. (2011). The march to reliable metrics: A half-century of coming closer to the truth. *Journal of Advertising Research, 51*(1), 124–135. https://doi.org/10.2501/JAR-51-1-124-135

Song, H., & Cho, J. (2021). Assessing (in)accuracy and biases in self-reported measures of exposure to disagreement: Evidence from linkage analysis using digital trace data. *Communication Methods and Measures, 15*(3), 190–210. https://doi.org/10.1080/19312458.2021.1935824

Southwell, B. G., Barmada, C. H., Hornik, R. C., & Maklan, D. M. (2002). Can we measure encoded exposure? Validation evidence from a national campaign. *Journal of Health Communication, 7*(5), 445–453. https://doi.org/10.1080/10810730290001800

Southwell, B. G., & Langteau, R. (2008). Age, memory changes, and the varying utility of recognition as a media effects pathway. *Communication Methods and Measures, 2*(1–2), 100–114. https://doi.org/10.1080/19312450802062380

Stone, A., & Shiffman, S. (1994). Ecological momentary assessment behavioral medicine. *Annals of Behavioral Medicine, 16*, 199–202. https://doi.org/10.1002/9781118453940.ch20

Stone, A., Shiffman, S. S., Atienza, A., & Nebeling, L. (Eds.). (2007). *The science of real-time data capture: Self-reports in health research*. Oxford University Press.

Sudman, S., & Bradburn, N. (1973). Effects of time and memory factors on response in surveys. *Journal of the American Statistical Association, 68*(344), 805–815. https://doi.org/10.1080/01621459.1973.10481428

Sudman, S., Finn, A., & Lannom, L. (1984). The use of bounded recall procedures in single interviews. *Public Opinion Quarterly*, *48*(2), 520–524. https://doi.org/10.1086/268847

Sullivan, H. W., Chen, W.-H., & Betts, K. R. (2021). Assessing the inclusion of foil items in a scale to measure recognition of health messages. *Communication Methods and Measures*, *15*(3), 211–221. https://doi.org/10.1080/19312458.2020.1768520

Thompson, C. P., Skowronski, J. J., Larsen, S. F., & Betz, A. L. (1996). *Autobiographical memory.* Erlbaum.

Thompson, C. P., Skowronski, J. J., & Lee, D. J. (1988). Telescoping in dating naturally occurring events. *Memory & Cognition*, *16*(5), 461–468. https://doi.org/10.3758/BF03214227

Thrasher, J., Huang, L., Perez, R., Niederdeppe, J., Santillan, E. A., & Alday, J. (2011). Evaluation of a social marketing campaign to support Mexico City's comprehensive smoke-free law. *American Journal of Public Health*, *101*(2), 328–335. https://doi.org/10.2105/AJPH.2009.189704

Timotijevic, L., Barnett, J., Shepherd, R., & Senior, V. (2009). Factors influencing self-report of mobile phone use: The role of response prompt, time reference and mobile phone use in recall. *Applied Cognitive Psychology*, *23*(5), 664–683. https://doi.org/10.1002/acp.1496

Tourangeau, R., Rips, L. J., & Rasinski, K. (2000). *The psychology of survey response.* Cambridge University Press.

Tversky, A., & Kahneman, D. (1973). Availability: A heuristic for judging frequency and probability. *Cognitive Psychology*, *5*(2), 207–232. https://doi.org/10.1016/0010-0285(73)90033-9

Tversky, A., & Koehler, D. J. (1994). Support theory: A nonextensional representation of subjective probability. *Psychological Review*, *101*(4), 547–567. https://doi.org/10.1037/0033-295X.101.4.547

Valkenburg, P. M., & Peter, J. (2013a). Five challenges for the future of media-effects research. *International Journal of Communication*, *7*, 197–215. http://ijoc.org/index.php/ijoc/article/download/1962/849

Valkenburg, P. M., Semetko, H. A., & de Vreese, C. H. (1999). The effects of news frames on readers' thoughts and recall. *Communication Research*, *26*(5), 550–569. https://doi.org/10.1177/009365099026005002

van der Voort, T. H. A., & Vooijs, M. W. (1990). Validity of children's direct estimates of time spent television viewing. *Journal of Broadcasting & Electronic Media*, *34*(1), 93–99. https://doi.org/10.1080/08838159009386729

van Hooff, M. L. M., Geurts, S. A. E., Kompier, M. A. J., & Taris, T. W. (2007). How fatigued do you currently feel? Convergent and discriminant validity of a single-item fatigue measure. *Journal of Occupational Health*, *49*(3), 224–234. https://doi.org/10.1539/joh.49.224

van Roekel, E., Keijsers, L., & Chung, J. M. (2019). A review of current ambulatory assessment studies in adolescent samples and practical recommendations. *Journal of Research on Adolescence*, *29*(3), 560–577. https://doi.org/10.1111/jora.12471

van Someren, M. W., Barnard, Y. F., & Sandberg, J. A. C. (1994). *The think aloud method: A practical guide to modelling cognitive processes*. Academic Press.

Vanden Abeele, M., Beullens, K., & Roe, K. (2013). Measuring mobile phone use: Gender, age and real usage level in relation to the accuracy and validity of self-reported mobile phone use. *Mobile Media & Communication*, *1*(2), 213–236. https://doi.org/10.1177/2050157913477095

Verbeij, T., Pouwels, J. L., Beyens, I., & Valkenburg, P. M. (2021). The accuracy and validity of self-reported social media use measures among adolescents. *Computers in Human Behavior Reports*, *3*, Article 100090. https://doi.org/10.1016/j.chbr.2021.100090

Verbeij, T., Pouwels, J. L., Beyens, I., & Valkenburg, P. M. (2022). Experience sampling self-reports of social media use have comparable predictive validity to digital trace measures. *Scientific Reports*, *12*(1), Article 7611. https://doi.org/10.1038/s41598-022-11510-3

Voorveld, H. A. M., van Noort, G., Muntinga, D. G., & Bronner, F. (2018). Engagement with social media and social media advertising: The differentiating role of platform type. *Journal of Advertising*, *47*(1), 38–54. https://doi.org/10.1080/00913367.2017.1405754

Vraga, E., Bode, L., & Troller-Renfree, S. (2016a). Beyond self-reports: Using eye tracking to measure topic and style differences in attention to social media content. *Communication Methods and Measures*, *10*(2–3), 149–164. https://doi.org/10.1080/19312458.2016.1150443

Vraga, E. K., Bode, L., Smithson, A., & Troller-Renfree, S. (2016b). Blurred lines: Defining social, news, and political posts on Facebook. *Journal of Information Technology & Politics, 13*(3), 272–294. https://doi.org/10.1080/19331681.2016.1160265

Watson, D., Clark, L. A., & Tellegen, A. (1988). Development and validation of brief measures of positive and negative affect: The PANAS scales. *Journal of Personality and Social Psychology*, *54*(6), 1063–1070. https://doi.org/10.1037/0022-3514.54.6.1063

Webster, J. G., Phalen, P. F., & Lichty, L. W. (2014). *Ratings analysis: Audience measurement and analytics* (4th ed.). Routledge.

Wenz, A., Keusch, F., & Bach, R. L. (2024). Measuring smartphone use: Survey versus digital behavioral data. *Social Science Computer Review*. Advance online publication. https://doi.org/10.1177/08944393231224540

Whipple, T. W., & McManamon, M. K. (1992). Primacy order effects in the measurement of trade magazine receipt and readership. *Journal of Advertising Research*, *32*(5), 24–29.

Woelke, J., & Pelzer, E. (2020). Cognitive assessment: Think-aloud and thought-listing technique. In J. Bulck (Ed.), *The international encyclopedia of media psychology*. Wiley. https://doi.org/10.1002/9781119011071.iemp0048

Wolfers, L. N., & Baumgartner, S. E. (2023, September 6–8). *Short but still valid: Validating single-item measures for key media psychology constructs for experience sampling research* [Paper presentation]. 13th Conference of the Media Psychology Division of the German Psychological Association (DGPs), Luxembourg.

Wonneberger, A., & Irazoqui, M. (2017). Explaining response errors of self-reported frequency and duration of TV exposure through individual and contextual factors. *Journalism & Mass Communication Quarterly*, *94*(1), 259–281. https://doi.org/10.1177/1077699016629372

Wrzus, C., & Neubauer, A. B. (2023). Ecological Momentary Assessment: A meta-analysis on designs, samples, and compliance across research fields. *Assessment*, *30*(3), 825–846. https://doi.org/10.1177/10731911211067538

Yzer, M. C., Vohs, K. D, Luciana, M., Cuthbert, B. N., & MacDonald III, A. W. (2011). Affective antecedents of the perceived effectiveness of antidrug advertisements: An analysis of adolescents' momentary and retrospective evaluations. *Prevention Science*, *12*(3), 278–288. https://doi.org/10.1007/s11121-011-0212-y

Zaichkowsky, J. L. (1985). Measuring the involvement construct. *Journal of Consumer Research*, *12*(3), 341–352. http://www.jstor.org/stable/254378

About the authors

Peter Neijens is an Honorary Fellow and Emeritus Professor at ASCoR. He held the Chair in Media and Persuasion at the University of Amsterdam until 2019. His research focuses on persuasive communication, in particular media use and effects in advertising and public opinion.

Theo Araujo is a Full Professor of Media, Organisations and Society, and Scientific Director of ASCoR. His research investigates the dynamic interplay between media and organizations, and what it means for society, with a special focus on trust and technology. His research has also a methodological focus on computational communication science and the implementation of large-scale data collection and analysis for communication research.

Judith Möller holds the chair of "Empirical Communication Research, especially Media Use and Social Media Effects" at the University of Hamburg and the Leibniz Institute for Media Research | Hans Bredow Institute (HBI).

Before 2023 she was an Associate Professor of Political Communication at the Amsterdam School of Communication Research.

Claes H. de Vreese is a Distinguished University Professor of AI and Society with a special emphasis on media and democracy. He co-directs the Gravitation Program AlgoSoc and the AI, Media and Democracy Lab. He holds the chair in Political Communication at the Amsterdam School of Communication Research.

5. Digital Trace Data

Abstract
In this chapter we discuss different methods for capturing digital tracking data for media exposure measurement, and discuss their advantages and disadvantages in section 5.1. We zoom in on data donation in section 5.2. Quality issues are discussed in section 5.3. The integration of digital trace data and self-report measures of media exposure is covered in section 5.4. Conclusions about digital trace data are drawn in section 5.5.

Keywords: platform centric, user centric, user tracking, data donation, quality issues

5.1 Capturing digital traces: different methods and their pros and cons

People's use of online media and digital platforms such as Google, Amazon, Apple, Facebook, TikTok or X leaves traces that form a wealth of data about how they are exposed to, attend to, and use media. These digital trace data, defined as "records of activity (trace data) undertaken through an online information system (thus, digital)" (Howison et al., 2011, p. 769), can be captured with different methods, known as tracking, registration, passive measurement, or passive monitoring. These methods enable researchers to obtain a variety of exposure metrics, ranging from the time using a specific application or visiting a website to the content that an individual may have been exposed to at a certain time. In addition, these methods also allow the capture of engagement metrics as well as other user activity (e.g., creation of content).

Table 5.1 provides an overview of different methods for recording digital traces of media behavior. An important distinction is between "site centric" (also "server centric", "platform centric", or "census data") methods versus "user centric" methods (Ohme et al., 2023; Webster et al., 2014; see also

Neijens, P., Araujo, T., Möller, J., & de Vreese, C. H. (2024). Digital trace data. In P. Neijens, T. Araujo, J. Möller, & C. H. de Vreese, *Measuring exposure and attention to media and communication. Solutions to wicked problems* (pp. 87-104). Amsterdam University Press. https://doi.org/10.5117/9789463723176_ch05

Chapter 2 in which this distinction was introduced). Site or platform centric methods rely on reporting or metrics captured directly by the website or platform. Website owners can, for example, make use of online services to record and analyze visits to their websites and how users navigate within their website (e.g., Google Analytics). Analytics data about a website or app is usually available only to the website owner, meaning that a researcher could only collect these data either by developing their own website (e.g., for an experiment with manipulated stimuli), or by entering a data sharing agreement with the owner of the website or service. Site or platform-centric methods are attractive as they allow measurement of media outlets with small audiences. However, they often lack background information about the individual user, and access to these data by (academic) researchers may be challenging.

In addition, depending on a particular platform's *Terms of Service*, the non-website owner may be able to extract data from the website — known as web scraping or web harvesting — using a web scraper (aka web crawler or bot), which is useful for extracting the content from the website. This information may be relevant for the researcher interested in enriching digital trace data collected through other user-centric methods discussed below (for example capturing the content of a specific website that a participant in a data donation study has visited), yet scraping by itself often does not necessarily provide relevant media exposure information.

Some digital platforms — for example X or YouTube — offer Application Programming Interfaces (APIs) through which the researcher can connect to the platform and collect data directly, for example the latest tweets about a topic, or the comments to a YouTube video. The conditions of APIs are subject to change. In the case of APIs, platforms — when they do provide an API for academic research — tend to mostly provide information about engagement (e.g., likes or shares) and very limited information on exposure (e.g., views). They usually just provide information about public content, typically at aggregated levels (e.g., how often a video was played in general, instead of to which pieces of content an *individual user* was exposed). A site centric method in physical spaces is the usage of grabbers or beacons to track the presence of mobile devices within or near a specific location.

The second type — user centric methods — tracks the media behavior of an individual user with the advantage that their media use can be combined with other characteristics that have been measured, for instance demographics, traits, interests, perceptions, and behavior via self-reports. We can distinguish several sub types. First, *registration of printed media exposure* with RFID technology that is able "to detect the openings and closings of

Table 5.1 Different methods to capture digital traces

Name	***What***	***How***
Site centric		
Site centric / server centric / platform centric / census data	Servers record website visits and how users navigate within a particular website.	Software installed in the server.
Web scraping / harvesting	Extracting data from websites.	Collecting data by web scrapers (web crawlers, bots).
APIs	Extracting data from digital platforms or services that offer an API.	Collecting the data by connecting to the API.
Beacons and (Wi-Fi or Bluetooth) grabbers	Captures the presence of a device (e.g., a smartphone) in a specific location. May track media exposure if combined with an object of interest (e.g., an outdoor advertising).	Beacons or grabbers are installed in a physical space, and track the presence of wireless devices.
User centric		
Passive measurement of printed media exposure	Openings and closings of printed magazines and the turning of pages within them.	RFID or related technology.
Household meter	Channel activity (registration of channel, on-demand video, DVDs, video games).	Set-top box.
People meter	In addition to the automatic monitoring of channel activity by the household meter, viewers indicate if they are watching/using the TV set, and when.	Set-top box plus handheld.
Router meter	Registers the IP traffic through the Wi-Fi network in the household.	Router software.
Mobile data technologies / Portable People Meters	Multimode data collection and capturing data in-the-moment.	Smartphone, wearables.
User tracking	Digital media use including date, time, and type of executed action.	Software application installed on user's device(s). Apps from the app store or tailor made installed on user's smartphone and other devices.
Data donation	Information that a digital platform or online service has on the individual (including their activity and potentially media exposure).	Data donation process, with the individual requesting their data from the relevant platform(s) and donating to academic research.

printed magazines and the turning of pages within them" (Mattlin & Gagen, 2013, p. 1). Second, the *household meter* which registers when the TV set (nowadays a display device for a variety of media) is on and the "device" (e.g., channel, on-demand video, DVDs, video games, content from the internet) to which it is tuned (Webster et al., 2014). The *people meter* is an extension of the household meter: in addition to the automatic monitoring of channel activity, respondents indicate when they are in the room, or — in some other countries — when they are in the room and watch the TV set. People meters are developed for audience measurement for advertisers, ad and media companies in syndicated research by companies such as Nielsen in the USA. The data is monitored for potential problems, for example through coincidental surveys, as compliance can be problematic due to factors such as "button-pushing fatigue" (Webster et al., 2014). Passive people meters have been developed that automatically recognize who is watching television. Another household meter is the *router meter* which registers the IP traffic through the Wi-Fi network in the home.

Mobile data technologies make use of wearables or smartphones with devices such as camera, microphone, GPS and scanner, which facilitate multimode data collection and capturing data in-the-moment (see Link et al., 2014). Early examples are Arbitron's Portable People Meter (PPM) and the Eurisko Media Monitor (EMM) which record an inaudible code that is embedded in the audio stream of audio and video programming, including broadcast TV, cable TV, radio, and audio/video content in stores, as well as audio-based commercials broadcast on these platforms (Fitzgerald, 2004; LaCour & Vavreck, 2014; Smit & Neijens, 2011; Taneja & Mamoria, 2012). An example in the field of public health is given by Lind et al. (2018) who built the Effortless Assessment of Risk States (EARS) tool: "The EARS tool captures multiple indices of a person's social and affective behavior via their naturalistic use of a smartphone. Although other mobile data collection tools exist, the EARS tool places a unique emphasis on capturing the content as well as the form of social communication on the phone. Signals collected include facial expressions, acoustic vocal quality, natural language use, physical activity, music choice, and geographical location" (p. 1).

User tracking through software applications or plug-ins installed on a user's device(s) is another method for collecting digital traces. Research apps — available from the app store or tailor-made — can be used to log app activity. These logging methods can capture a wealth of activities related to exposure metrics such as internet browsing, radio listening, tv viewing, gaming, calling, texting, etc. The data contains information on most used apps, day, time and length of app use, screen on/off, calls in/out, SMS sent/

received, passive reading, typing, photo taking, and spontaneous use versus response-based actions, to name a few (Araujo et al., 2017; Boase & Ling, 2013; Christner et al., 2022; Ellis, 2019; Geyer et al., 2022; Rauwers et al., 2020; Ryding & Kuss, 2020; Verbeij et al., 2021, 2022). In more advanced formats of user tracking, such as in the study of Ram et al. (2019) (four respondents), automated smartphone screenshots were obtained every five seconds that the device was activated providing a detailed and granular view of one's usage of the smartphone and, as a consequence, of one's media exposure.

Christner et al. (2022) provide a critical review and classification of automated tracking approaches for studying online media exposure, apt to guide research decisions on the appropriate tracking approach and tool.

5.2 Data donation

A different approach to collect digital exposure data is by asking participants to donate their digital traces, and from these digital traces extract metrics associated with media exposure. This can be done using different processes to gather the data. Some of the most common methods are:

1. Reuse functions integrated into the operating systems of devices such as smartphones that are intended to give users insight and control over their smartphone or app usage (e.g., reports of screen time or battery usage), with researchers asking participants to donate a screenshot or video of these reports to academic research (Baumgartner et al., 2023; Ohme et al., 2021).
2. Use browser plugins to, for example, retrieve the web history of individuals who are willing to install the plugin and donate their data to the researcher (e.g., Menchen-Trevino, 2016), and as such capture individual exposure to different websites, pages or even news stories.
3. Ask participants to request a copy of the data that digital platforms have on them, and donate a (subset of) these data. Legislation (e.g., EU's General Data Protection Regulation (GDPR)) requires companies to make digital trace data available to their users that request them. "Data Download Packages (DDPs) may include usage information (e.g., login history), activity (e.g., posts created, messages sent), and profiling done by the platform (e.g., inferred interests or categories), among other information" (Araujo et al., 2022, p. 375). Among these data, researchers can capture traces related to exposure (e.g., the list of videos watched by a participant on TikTok or YouTube, or the last items viewed on Facebook). Researchers can collaborate with users who are willing to

donate these data, for example, their Instagram data (van Driel et al., 2022) or Facebook data (Thorson et al., 2019). Several research groups have worked to create open-source frameworks to facilitate the donation process (e.g., Araujo et al., 2022; Boeschoten, Ausloos, et al., 2022; Boeschoten, Mendrik, et al., 2022).

The usage of data donation methods has a number of advantages. When discussing the example of DDPs (option 3, above), van Driel and colleagues (2022), highlight: "First, DDPs provide a full overview of the uses of a platform regardless of whether it was accessed via the phone, tablet, or laptop. Second, DDPs capture all user interactions with the platform from the moment a user created an account until the moment of the download request. Third, because data of platform users is collected automatically by the social media companies, no (smartphone) applications need to be installed and thus researcher bias is limited. Fourth, all information is timestamped and separated into text and media files, categorized per social media activity" (p. 267).

We give some examples of data donation studies. Wojcieszak et al. (2022) used Web Historian (an open-source extension for Google Chrome that accesses respondents' browser history stored on their computers) to collect the browser history of participants allowing them to determine exposure to news websites by analyzing the URLs that participants had visited. Data in the study of Baumgartner et al. (2023) were collected with the iOS Battery Section that is more granular as it "displays which apps are used for how many minutes each hour of the day, including information about on screen as well as background usage of these apps screen shots" (p. 2). Ohme et al.'s (2021) study made use of iOS 12's Screen Time feature which provides information on screen time and number of pickups, as well as most-used apps and websites. In the study of Thorson et al. (2019) respondents were asked to download their own Facebook DDP and upload relevant files to the researchers: an "index" file, which lists pages liked by the participant, and an "'ads' file, which provides 'A list of topics on which you may be targeted based on your stated likes, interests and other data you put in your Timeline' (Facebook Help Center, n.d.)" (p. 188).

5.3 Quality

The advantages of digital trace data are clear: unobtrusive, high precision and granularity. On the other hand: retrieving, storing, and processing of the (often large volume of) data involve high cost, there are (potential) conceptual, ethical, and technical issues, and (biased) non-response can be substantial (e.g., Araujo et al., 2017; Boeschoten, Ausloos, et al., 2022; Boeschoten, Mendrik, et al., 2022; Deng et al., 2019; Jurgens et al., 2019; Ryding & Kuss, 2020; Scharkow, 2016; Stier et al., 2020; van Driel et al., 2022; Verbeij et al., 2021, 2022; Webster et al., 2014). These quality issues can be derived from the "Total Error Framework for Digital Traces of Human Behavior on Online Platforms (TED-On)," see chapter 3 (Sen et al., 2021).

Conceptual issues

First conceptual issues: What theoretical concept do digital traces measure? What level of exposure is measured with audio recordings, page opens, or likes? To what extent do users engage with the message they are exposed to? As the answers are not obvious without further information, studies should make clear what the meaning of their measures is.

Another issue is that digital trace data are often "found data," not based on research designs with carefully developed measurement instruments. "Through this, the field runs the risk of producing largely data-driven studies instead of testing carefully developed hypotheses based on current theoretical debates" (Jungherr, 2019, p. 11).

Data issues

Digital data are unstructured. They may include website visits, liking of messages, comments on YouTube videos, photos and videos. Automatic coding of texts including machine learning-based procedures is used to code the content of these communications, but that may come with serious measurement errors (for example in case of use of ironic language; Lewis et al., 2013; Mostafa, 2013).

Ethical issues

There are ethical questions about privacy and surveillance. Capturing digital traces of individuals requires respecting Terms and Services of the platforms and informed consent from the people whose data is collected. The latter is not always possible, such as when the data contains information about "other users" who could not, or were not, asked for their consent (Al Baghal et al., 2019; Stier et al., 2020). "Inclusiveness" is also an ethical issue as some respondents lack the necessary skills to participate (Baumgartner et al., 2023).

Technical issues

The list of technical issues is long, and highly dependent on the method used. For instance, different operating systems (iOS versus Android) do not give access to the same data, some research apps run on the one operating system and not on the other. Other issues include crashes, bugs in software and hardware, and applications that may run in the background (thus apps are active while users do other things). Also, clicking on a page does not necessarily mean that the user has viewed all information on that page (on laptop and tablet screens scrolling is necessary as the screen size is too small to display an entire page).

Van Driel et al. (2022) extensively discuss issues related to data donation: the structure of research apps is not uniform across respondents and time, data labels (the variables) are not always clear, coding of the data is time-consuming and complicated, and device space and quality of the internet connection required to upload the data may be an issue for some respondents.

Lack of transparency

Methods that collect digital traces are dependent on how these digital traces are created in the first place, which, often, relies on processes that are not always transparent to researchers. For example, measures obtained via scraping, APIs or data donation may provide number of views or engagement for a specific content (e.g., a video, a website, or a tweet) yet these measures may not be evidence of *only* human activity. Automated scripts (e.g., bots) which "visit" websites can artificially increase the number of visits, or follow accounts and/or engage automatically with content on social media. While some of this activity may be part of how websites operate (e.g., bots from search engines to index online content), a share may be related to fraud, for instance in marketing applications where advertisers "pay per click," or in case of bots on X automatically disseminating misinformation and inflating retweet counts. In addition, data collected via these methods is also often produced "under the algorithm" (Wagner et al., 2021), meaning that algorithms (e.g., recommender systems) and feedback loops may influence the data that researchers use as the basis for their analyses.

Another challenge has to do with the clarity on what the measures mean or how they are operationalized. For example, documentation on the data and measures is often missing (e.g., for data donation, van Driel et al., 2022), or aggregated by platforms without much information on their logics (e.g., pre-defined categories of apps as seen in mobile data donations, Ohme et al., 2021).

Measurement issues

Cernat et al. (2024) estimated the measurement quality of survey and digital trace data on smartphone usage with a MultiTrait MultiMethod (MTMM) model: "The experimental design included five topics relating to the use of smartphones (traits) measured with five methods: three different survey scales (a 5- and a 7-point frequency scale and an open-ended question on duration) and two measures from digital trace data (frequency and duration). We show that surveys and digital trace data measures have very low correlation with each other. We also show that all measures are far from perfect and, while digital trace data appears to have often better quality compared to surveys, that is not always the case" (p. 1).

Coverage

User tracking methods often do not include all devices of a respondent (e.g., home computer, work computer, tablet, smartphone), and as such do not provide a complete picture of a user's digital life. Another issue is that the same devices can also be used by another person, which can also lead to incorrect conclusions. A limitation of data donation data is that it may be limited to a specific platform (e.g., Facebook) and lack activities that individuals do when they leave such a platform.

Non-response and biased sample issues

If researchers depend for their digital trace data on data access policies set by service providers "*n* = all becomes *n* = sample with unknown properties from unknown populations determined by third parties" (Jungherr, 2019, p. 11). When researchers depend on users' willingness to participate in surveys that record their media behavior, the non-response problem comes into play. There are several sources of non-response and (potentially) biased samples: non-response in the prior study, non-consent to the tracking part of the study, and non-response to the tracking study (Stier et al., 2020).

Research shows that response in digital trace studies is usually below 35% and substantively biased (Stier et al., 2020). Response rates differ depending on population, topic, sampling strategy, length of data collection period, incentive, sensitivity of the data that is collected, and technical complexity.

For instance, Keusch et al. (2019) studied factors that influenced smartphone users' willingness to participate in passive mobile data collection by a research app on the participant's smartphone (hypothetical vignette study; n = 1,947).They found that only 35% of their respondents (online panel in Germany who regularly participate in surveys) "indicated their willingness to download an app that would passively collect technical characteristics of the phone, cell network parameters, geographic location, app usage, and browser history, as well as the number of incoming and outgoing phone calls and text messages. About a third (39 percent) of respondents were not willing to download an app under any of the conditions described in the vignettes" (p. 229). They also found that participation "is strongly influenced by the incentive promised for study participation but also by other study characteristics (sponsor, duration of data collection period, option to switch off the app) as well as respondent characteristics (privacy and security concerns, smartphone experience)" (p. 211).

Another phenomenon that may add to these representative issues is that respondents might adapt their behavior knowing that their behavior is recorded. Also: Internet usage is usually logged for a limited time, so the measurement depends on the period chosen for the survey.

Data donation issues

Specifically, for data donation studies, Boeschoten, Ausloos, and colleagues (2022) proposed an error framework based on the total error framework (see Chapter 3), which summarizes some of the potential measurement and sampling issues that may be applicable to digital trace data, as well as some recommendations (see Table 5.2).

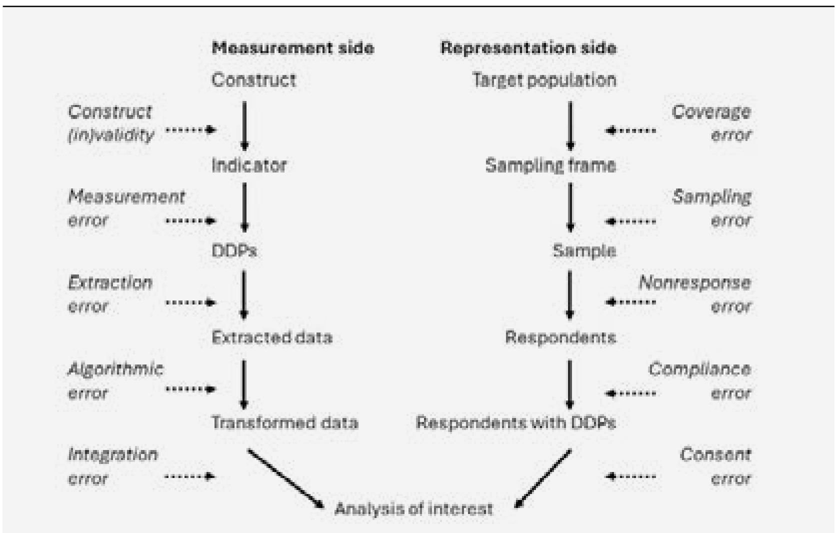
On the measurement side, researchers should address the potential for:

- Construct (in)validity and measurement errors: Researchers should consider triangulating the proposed measurement with other measures (e.g., self-reports in a survey, different indicators within the same digital trace data collection strategy).
- Extraction errors: Given the complex nature of the digital trace data being collected — which is often also unstructured — researchers should carefully assess the process of extracting the relevant measures from the digital traces. This can be done by extensive testing, on the one hand, and triangulation, on the other.
- Algorithmic errors: The extraction of measures from digital trace data often not only requires actual data extraction (i.e., finding a specific measure — e.g., screen time — from digital traces), but also requires

categorization and aggregation of the measure (e.g., categorizing log-like data — e.g., a set of timestamped rows with app names — into a meaningful measure — e.g., exposure time per type of app). This process often relies on automated classifications (or the training of supervised machine learning models), and thus researchers should inspect the reliability of this categorization process, as outlined in Chapter 3.

On the representation side, Boeschoten, Ausloos, et al. (2022) suggest that researchers should be mindful that digital trace data collection methods, given their potential complexity, require special attention to the same types of errors as those seen in surveys, i.e., coverage, sampling, nonresponse, compliance, or consent errors (see Chapter 3). In all of these, given the nature of these methods — which for example require respondents to install a plugin (in the case of tracking), or request and upload their data (in the case of data donation) — researchers should be careful with incomplete or missing data due to non-random factors. For example, participants with higher levels of privacy concerns may be less prone to accept to participate in one of these studies (thus leading to nonresponse errors or bias) or those with lower levels of technical expertise being less likely to install a plugin or donate their data (compliance errors or bias).

Table 5.2 Error framework for data donation



Source: Boeschoten, Ausloos, et al. (2022). "Total error framework" for social-scientific data collection with DDPs based on Amaya et al. (2020).

Response in data donation studies

We illustrate the (non-) response issue in data donation studies in Table 5.3.

Table 5.3 Response and compliance in data donation studies: two examples

Example 1 Mobile data donation study of Ohme et al. (2021)

- Representative sample of Dutch speaking iPhone users in the Netherlands (*n* = 404 participants, (AAPOR 1) response rate = 54%).
- Of the 404 respondents, 307 (75.8% collecting sample) agreed at the end of the first wave of the survey to be contacted again and confirmed they had responded positively to the request to turn on the Screen Time function on their phone.
- 122 respondents (retention rate 40%) actually finished the second wave of the survey.
- Of those 122 respondents, 47 shared their mobile log data successfully with the researchers by uploading their screenshots, presenting the final donator sample.
- The other 65 respondents either did not start an upload, uploaded incomplete or bogus content, or did not enter the correct information (e.g., their unique identifier) in the process of switching from taking the survey on a computer to uploading the screenshots on a smartphone.
- Overall, 11.6% of participants of the full sample ultimately donated their mobile log data to this study.
- It seems that mobile privacy literacy and not perceived risks per se affected the decision to donate data. Overall, sample biases seem rather marginal.

Example 2 Instagram data donation study (van Driel et al., 2022)

- 745 potential participants (adolescents, all 8th and 9th graders, recruited via a large secondary school in the Netherlands).
- 400 participants received parental consent (53.7%).
- 388 provided assent for the larger study (52.1%).
- 287 obtained parental consent for the data download portion of the study (38.5%).
- 209 indicated to have an Instagram account.
- 148 provided informed assent and automatically proceeded with the data donation process.
- 102 participants provided useable Instagram DDPs (48.8% of the 209 participants with an Instagram account and parental consent for the data download part of the study); 1.37% of the 745 potential study participants).
- The sample of participants who donated their DDPs was comparable to the sample who did not donate in terms of age, but the DDP sample consisted of more girls and fewer participants who followed a lower educational track.

Source: Ohme et al. (2021); van Driel et al. (2022).

Factors that influence user's willingness to donate data are studied by Pfiffner and Friemel (2023). Based on a national (Swiss) online survey (*n* = 833), the effects of three types of variables were found: platform type (higher willingness to donate YouTube data compared to Facebook, Instagram, or Google), data type (higher willingness for lower perceived sensitivity data and higher perceived relevance of the data), and individual level factors (more favorable attitudes toward data donation and the donation purpose, lower privacy concerns, higher perceived behavioral control to request and submit the data).

5.4 Integrating digital trace data with self-reports

Combining digital trace data and self-reports, at the aggregate or individual level, ex-ante or ex-post (Stier et al., 2020), is useful for cross-validation and improvement of media exposure measurements. In Chapter 4 we discussed Prior's (2009b) study that compared self-reports of media exposure with Nielsen data on TV viewing (an example of linking digital trace data with self-reports at the *aggregate* level, ex-post). The studies that compared self-reports with *individual* trace data are numerous (see Chapter 4, for instance Araujo et al., 2017; Parry et al., 2021; Verbeij et al., 2021, 2022; Wonneberger & Irazoqui, 2017).

Linking self-report measures and digital trace data can also be useful for improvement of substantial analysis, for instance by measuring online news use with digital trace data, and by measuring offline news use and other variables of interest in a survey (e.g., Burke & Kraut, 2016; Kristensen et al., 2017; Möller et al., 2020; Stier et al., 2020).

An interesting example of integrating digital trace data with interview data is the "stimulated recall" study by Griffioen et al. (2020a, 2020b). They collected (uninformed) participants' social media behavior through video footage and in-phone data while they were waiting for their interview. In the interview these data were reviewed with the participants discussing their (reasons) for use.

5.5 Concluding remarks

Despite their issues, digital trace methods are an invaluable addition to the researcher's toolkit. However, there is an urgent need for further research into the conceptual, ethical, technical, and response aspects discussed in this chapter, as well as broader discussions about the role that digital platforms often play as gatekeepers of this information. The relationship between digital trace data with self-reports, and the role of individual, media, use, situation, and other factors in this, are high on the research agenda. These are not only methodological questions (which measure is best under which circumstances?), but also substantive issues, because they provide insight into how the frequency and duration of "physical" confrontations with media are experienced and processed by the user.

References

Al Baghal, T., Sloan, L., Jessop, C., Williams, M., & Burnap, P. (2020). Linking Twitter and survey data: The impact of survey mode and demographics on consent rates across three UK studies. *Social Science Computer Review*, *38*(5), 517–532. https://doi.org/10.1177/0894439319828011

Amaya, A., Biemer, P. P., & Kinyon, D. (2020). Total error in a big data world: Adapting the TSE framework to big data. *Journal of Survey Statistics and Methodology*, *8*(1), 89–119. https://doi.org/10.1093/jssam/smz056

Araujo, T., Ausloos, J., van Atteveldt, W., Loecherbach, F., Moeller, J., Ohme, J., Trilling, D., van de Velde, B., de Vreese, C., & Welbers, K. (2022). OSD2F: An open-source data donation framework. *Computational Communication Research*, *4*(2), 372–387. https://doi.org/10.5117/CCR2022.2.001.ARAU

Araujo, T., Wonneberger, A., Neijens, P., & de Vreese, C. (2017). How much time do you spend online? Understanding and improving the accuracy of self-reported measures of internet use. *Communication Methods & Measures*, *11*(3), 173–190. https://doi.org/10.1080/19312458.2017.1317337

Baumgartner, S. E., Sumter, S. R., Vladislav, P., & Wiradhany, W. (2023). A novel iOS data donation approach: Automatic processing, compliance and reactivity in a longitudinal study. *Social Science Computer Review*, *41*(4), 1456–1472. https://doi.org/10.1177/08944393211071068

Boase, J., & Ling, R. S. (2013). Measuring mobile phone use: Self-report versus log data. *Journal of Computer-Mediated Communication*, *18*(4), 508–519. https://doi.org/10.1111/jcc4.12021

Boeschoten, L., Ausloos, J., Möller, J. E., Araujo, T., & Oberski, D. L. (2022). A framework for privacy preserving digital trace data collection through data donation. *Computational Communication Research*, *4*(2), 388–423. https://doi.org/10.5117/CCr2022.2.002.BOES

Boeschoten, L., Mendrik, A., van der Veen, E., Vloothuis, J., Hu, H., Voorvaart, R., & Oberski, D. L. (2022). Privacy-preserving local analysis of digital trace data: A proof-of-concept. *Patterns*, *3*(3), Article 100444. https://doi.org/10.1016/j.patter.2022.100444

Burke, M., & Kraut, R. E. (2016). The relationship between Facebook use and well-being depends on communication type and tie strength. *Journal of Computer-Mediated Communication*, *21*(4), 265–281. https://doi.org/10.1111/jcc4.12162

Cernat, A., Keusch, F., Bach, R. L., & Pankowska, P. K. (2024). Estimating measurement quality in digital trace data and surveys using the MultiTrait MultiMethod Model. Social Science Computer Review, 0(0). https://doi.org/10.1177/08944393241254464

Christner, C., Urman, A., Adam, S., & Maier, M. (2022). Automated tracking approaches for studying online media use: A critical review and recommendations. *Communication Methods and Measures*, *16*(2), 79–95. https://doi.org/10.1080/19312458.2021.1907841

Deng, T., Kanthawala, S., Meng, J., Peng, W., Kononova, A., Hao, Q., Zhang, Q., & Prabu, D. (2019). Measuring smartphone usage and task switching with log tracking and self-reports. *Mobile Media & Communication*, *7*(1), 3–23. https://doi.org/10.1177/2050157918761491

Ellis, D. A. (2019). Are smartphones really that bad? Improving the psychological measurement of technology-related behaviors. *Computers in Human Behavior*, *97*, 60–66. https://doi.org/10.1016/j.chb.2019.03.006

Facebook Help Center. (n.d.). *What categories of my Facebook data are available to me?* https://www.facebook.com/help/930396167085762

Fitzgerald, J. (2004). Evaluating return on investment of multimedia advertising with a single-source panel: A retail case study. *Journal of Advertisement Research*, *44*(3), 262–270. https://doi.org/10.1017/S0021849904040309

Geyer, K., Ellis, D.A., Shaw, H., & Davidson, B. I. (2022). Open-source smartphone app and tools for measuring, quantifying, and visualizing technology use. *Behavior Research Methods*, *54*, 1–12. https://doi.org/10.3758/s13428-021-01585-7

Griffioen, N., van Rooij, M. M. J. W., Lichtwarck-Aschoff, A., & Granic, I. (2020a). A stimulated recall method for the improved assessment of quantity and quality of social media use. *Journal of Medical Internet Research*, *22*(1), Article e15529. https://doi.org/10.2196/15529

Griffioen, N., van Rooij, M. M. J. W., Lichtwarck-Aschoff, A., & Granic, I. (2020b). Toward improved methods in social media research. *Technology, Mind, and Behavior*, *1*(1). https://doi.org/10.1037/tmb0000005

Howison, J., Wiggins, A., & Crowston, K. G. (2011). Validity issues in the use of social network analysis with digital trace data. *Journal of the Association of Information Systems*, *12*(12), 767–797. https://doi.org/10.17705/1jais.00282

Jungherr, A. (2019). Normalizing digital trace data. In N. J. Stroud & S. McGregor (Eds.), *Digital discussions: How big data informs political communication* (pp. 9–35). Routledge. https://doi.org/10.4324/9781351209434-2

Jurgens, P., Stark, B., & Magin, M. (2019). Two half-truths make a whole? On bias in self-reports and tracking data. *Social Science Computer Review*, *38*(5), 600–615. https://doi.org/10.1177/0894439319831643

Keusch, F., Struminskaya, B., Antoun, C., Couper, M. P., & Kreuter, F. (2019). Willingness to participate in passive mobile data collection. *Public Opinion Quarterly*, *83*(1), 210–235. https://doi.org/10.1093/poq/nfz007

Kristensen, J. B., Albrechtsen, T., Dahl-Nielsen, E., Jensen, M., Skovrind, M., & Bornakke, T. (2017). Parsimonious data: How a single Facebook like predicts voting behavior in multiparty systems. *PLOS ONE*, *12*(9), Article e0184562. https://doi.org/10.1371/journal.pone.0184562

LaCour, M. J., & Vavreck, L. (2014). Improving media measurement: Evidence from the field. *Political Communication*, *31*(3), 408–420. https://doi.org/10.1080/10584609.2014.921258

Lewis, S. C., Zamith, R., & Hermida, A. (2013). Content analysis in an era of Big Data: A hybrid approach to computational and manual methods. *Journal of Broadcasting and Electronic Media*, *57*(1), 34–52. https://doi.org/10.1080/08838151.2012.761702

Lind, M. N., Byrne, M. L., Wicks, G., Smidt, A. M., & Allen, N. B. (2018). The effortless assessment of risk states (EARS) tool: An interpersonal approach to mobile sensing. *JMIR Mental Health*, *5*(3), Article e10334. https://doi.org/10.2196/10334

Link, M. W., Murphy, J., Schober, M. F., Buskirk, T. D., Childs, J. H., & Tesfaye, C. L. (2014). Mobile technologies for conducting, augmenting and potentially replacing surveys: Report of the AAPOR task force on emerging technologies in public opinion research. *Public Opinion Quarterly*, *78*(4), 779–787. https://doi.org/10.1093/poq/nfu054

Mattlin, J. A., & Gagen, J. (2013, October 13–16). *Passive measurement of print revisited: The case of digital magazines* [Paper presentation]. Print and Digital Research Forum, Nice, France.

Menchen-Trevino, E. (2016). Web historian: Enabling multi-method and independent research with real-world web browsing history data. *Proceedings of IConference 2016, USA*. https://doi.org/10.9776/16611

Möller, J., van de Velde, R. N., Merten, L., & Puschmann, C. (2020). Explaining online news engagement based on browsing behavior: Creatures of habit?. *Social Science Computer Review*, *38*(5), 616–632. https://doi.org/10.1177/0894439319828012

Mostafa, M. M. (2013). More than words: Social networks' text mining for consumer brand sentiments. *Expert Systems with Applications*, *40*(10), 4241–4251. https://doi.org/10.1016/j.eswa.2013.01.019

Ohme, J., Araujo, T., Boeschoten, L., Freelon, D., Ram, N., Reeves, B. B., & Robinson, T. N. (2023). Digital trace data collection for social media effects research: APIs, data donation, and (screen) tracking. *Communication Methods and Measures*. Advance online publication. https://doi.org/10.1080/19312458.2023.2181319

Ohme, J., Araujo, T., de Vreese, C. H., & Piotrowski, J. T. (2021). Mobile data donations: Assessing self-report accuracy and sample biases with the iOS Screen Time function. *Mobile Media & Communication*, *9*(2), 293–313. https://doi.org/10.1177/2050157920959106

Parry, D., Davidson, B. I., Sewall, C., Fisher, J. T., Mieczkowski, H., & Quintana, D. S. (2021). A systematic review and meta-analysis of discrepancies between logged and self-reported digital media use. *Nature Human Behavior*, *5*(11), 1535–1547. https://doi.org/10.1038/s41562-021-01117-5

Pfiffner, N., & Friemel, T. N. (2023). Leveraging data donations for communication research: Exploring drivers behind the willingness to donate. *Communication Methods and Measures*, *17*(3), 227–249. https://doi.org/10.1080/19312458.2023.2176474

Prior, M. (2009b). The immensely inflated news audience: Assessing bias in self-reported news exposure. *Public Opinion Quarterly*, *73*(1), 130–143. https://doi.org/10.1093/poq/nfp002

Ram, N., Yang, X., Cho, M.-J., Brinberg, M., Muirhead, F., Reeves, B., & Robinson, T. N. (2019). Screenomics: A new approach for observing and studying individuals' digital lives. *Journal of Adolescent Research*, *35*(1), 16–50. https://doi.org/10.1177/0743558419883362

Rauwers, F., Voorveld, H. A. M., & Neijens, P. C. (2020). Explaining perceived interactivity effects on attitudinal responses: A field experiment on the impact of external and internal communication features in digital magazines. *Journal of Media Psychology*, *32*(3), 130–142. https://doi.org/10.1027/1864-1105/a000266

Ryding, F. C., & Kuss, D. J. (2020). Passive objective measures in the assessment of problematic smartphone use: A systematic review. *Addictive Behaviors Reports*, *11*, Article 100257. https://doi.org/10.1016/j.abrep.2020.100257

Scharkow, M. (2016). The accuracy of self-reported internet use—A validation study using client log data. *Communication Methods and Measures*, *10*(1), 13–27. https://doi.org/10.1080/19312458.2015.1118446

Sen, I., Flöck, F., Weller, K., Weiß, B., & Wagner, C. (2021). A total error framework for digital traces of human behavior on online platforms. *Public Opinion Quarterly*, *85*(S1), 399–422. https://doi.org/10.1093/poq/nfab018

Smit, E. G., & Neijens, P. C. (2011). The march to reliable metrics: A half-century of coming closer to the truth. *Journal of Advertising Research*, *51*(1), 124–135. https://doi.org/10.2501/JAR-51-1-124-135

Stier, S., Breuer, J., Siegers, P., & Thorson, K. (2020). Integrating survey data and digital trace data: Key issues in developing an emerging field. *Social Science Computer Review*, *38*(5), 503–516. https://doi.org/10.1177/0894439319843669

Taneja, H., & Mamoria, U. (2012). Measuring media use across platforms: Evolving audience information systems. *The International Journal on Media Management*, *14*(2), 121–140. https://doi.org/10.1080/14241277.2011.648468

Thorson, K., Cotter, K., Medeiros, M., & Pak, C. (2019). Algorithmic inference, political interest, and exposure to news and politics on Facebook. *Information, Communication & Society*, *24*(2), 183–200. https://doi.org/10.1080/1369118X.2019.1642934

van Driel, I. I., Giachanou, A., Pouwels, J., Boeschoten, L., Beyens, I., & Valkenburg, P. M. (2022). Promises and pitfalls of social media data donations. *Communication Methods and Measures*, *16*(4), 266–282. https://doi.org/10.1080/19312458.2022.2109608

Verbeij, T., Pouwels, J. L., Beyens, I., & Valkenburg, P. M. (2021). The accuracy and validity of self-reported social media use measures among adolescents. *Computers in Human Behavior Reports*, *3*, Article 100090. https://doi.org/10.1016/j.chbr.2021.100090

Verbeij, T., Pouwels, J. L., Beyens, I., & Valkenburg, P. M. (2022). Experience sampling self-reports of social media use have comparable predictive validity to digital trace measures. *Scientific Reports*, *12*(1), Article 7611. https://doi.org/10.1038/s41598-022-11510-3

Wagner, C., Strohmaier, M., Olteanu, A., Kıcıman, E., Contractor, N., & Eliassi-Rad, T. (2021). Measuring algorithmically infused societies. *Nature*, *595*(7866), 197–204. https://doi.org/10.1038/s41586-021-03666-1

Webster, J. G., Phalen, P. F., & Lichty, L. W. (2014). *Ratings analysis: Audience measurement and analytics* (4th ed.). Routledge.

Wojcieszak, M., Menchen-Trevino, E., Goncalves, J. F. F., & Weeks, B. (2022). Avenues to news and diverse news exposure online: Comparing direct navigation, social media, news aggregators, search queries, and article hyperlinks. *The International Journal of Press/Politics*, *27*(4), 860–886. https://doi.org/10.1177/19401612211009160

Wonneberger, A., & Irazoqui, M. (2017). Explaining response errors of self-reported frequency and duration of TV exposure through individual and contextual factors. *Journalism & Mass Communication Quarterly*, *94*(1), 259–281. https://doi.org/10.1177/1077699016629372

About the authors

Peter Neijens is an Honorary Fellow and Emeritus Professor at ASCoR. He held the Chair in Media and Persuasion at the University of Amsterdam until 2019. His research focuses on persuasive communication, in particular media use and effects in advertising and public opinion.

Theo Araujo is a Full Professor of Media, Organisations and Society, and Scientific Director of ASCoR. His research investigates the dynamic interplay between media and organizations, and what it means for society, with a special focus on trust and technology. His research has also a methodological focus on computational communication science and the implementation of large-scale data collection and analysis for communication research.

Judith Möller holds the chair of "Empirical Communication Research, especially Media Use and Social Media Effects" at the University of Hamburg and the Leibniz Institute for Media Research | Hans Bredow Institute (HBI). Before 2023 she was an Associate Professor of Political Communication at the Amsterdam School of Communication Research.

Claes H. de Vreese is a Distinguished University Professor of AI and Society with a special emphasis on media and democracy. He co-directs the Gravitation Program AlgoSoc and the AI, Media and Democracy Lab. He holds the chair in Political Communication at the Amsterdam School of Communication Research.

6. Observation

Abstract
Observation of media behavior, both in the lab and at home, is a well-known research method in media and communication research. In recent years, the focus of observational methods has shifted to online observation of behavior (e.g., Harrington, 2020). We discuss different types of observation studies and their pros and cons in section 6.1. Concluding remarks can be found in section 6.2.

Keywords: observer types, observation protocols, digital ethnography, quality issues

6.1 Different methods and their pros and cons

Observation of media behavior is applied in ethnographic studies (e.g., Lull, 1982, 1988), research into responses to commercials and commercial breaks (e.g., Krugman et al., 1995; Steiner, 1966), and studies into children and media (e.g., Buijzen et al., 2008). Observation can take different forms. "A distinction can be made between an intervening versus non-intervening observer (thus in naturally occurring settings); between participant versus nonparticipant observation; between visible (overt) versus nonvisible participation; between observation with versus without knowledge of the subject; and between observations with or without (hidden) hardware (e.g., microphone, camera)" (Araujo & Neijens, 2020, p. 3). For example, observers of the TV viewer usually document different behaviors such as "eyes on the screen" or "attention level" using a codebook. These studies are often supplemented with field notes (in which the observer reports the context of the behavior), and interviews or surveys. Field notes, photos, videos and the like are content analyzed afterwards.

Neijens, P., Araujo, T., Möller, J., & de Vreese, C. H. (2024). Observation. In P. Neijens, T. Araujo, J. Möller, & C. H. de Vreese, *Measuring exposure and attention to media and communication. Solutions to wicked problems* (pp. 105-111). Amsterdam University Press. https://doi.org/10.5117/9789463723176_ch06

Steiner was one of the first who studied viewer behavior with the "written records of trained and hopefully unobserved observers" … [he conducted] "a spy study in which one member of each family surreptitiously observed and recorded the viewing behavior of another member" (Steiner, 1966, pp. 272–273). The general instructions for his observers are listed in Table 6.1. His research showed that only 183 of the original 325 observers (who made a total of 47,823 observations) survived the various reliability and quality checks consisting of internal consistency checks and a comparison with interview data. This demonstrates that observing behavior is not an easy task and that the data obtained is not necessarily of high quality. Positive was that only very few of the 183 remaining observers reported that their subject was suspicious of their observer task.

Table 6.1 Instructions to the observers in the study of Steiner (1966)

"Observers were first to memorize the four rating scales used and then, whenever possible, simply lounge in the same room with the viewing subject and record his behavior during commercials and other NPE's [Non Program Elements] on unlabeled recording sheets. Hopefully, the nature of their observations was to go entirely undetected by the subject. Toward that end, we provided a standard cover story: If questioned, they were to say something general about doing a school assignment; and, if pressed for specifics, to explain that it dealt with recording their own reactions to various segments of television programming. This was to explain their presence and activity in front of the television set, to take the emphasis off the commercials, and to make it appear natural to make ratings during the commercial breaks.
…
The observers coded behavior, e.g., during a commercial break, based on the following coding scheme:
1. Full attention. Stays in (chair) and watches all or almost all (attention to visual and audio).
2. Partial attention. Stays in (chair) but does not pay (full) attention (turns around, talks, etc.) (exposure to visual and audio).
3. Gets up but stays in room (gets something, make phone call in room, etc.).
4. Leaves the room.
5. Not in room at onset."

Source: Steiner (1966, pp. 273–274).

Another observational study was conducted by Krugman et al. (1995): sixteen trained observers (grad students) collected data on sixty-four subjects — members or visitors in the households. They registered — for one hour — eyes-on-screen time (using stopwatches and coding sheets) and recorded activities that occur during television viewing. Observer field notes were added to the study. In the interviews, the respondents were asked to indicate their perception of their TV behavior and memory of the media content (i.e., the commercials). What is also interesting about the findings in this study is that subjects with no eyes on the screen could visualize (and remember) the content of TV commercials just by hearing them.

In Lull's (1982) study, observers spent no less than two days with the families randomly assigned to them and returned a third day to interview each family member. "They ate dinner with the families and generally took part in all their activities. To the degree it was possible, families were asked to ignore the presence of the observer and carry out their routines in normal fashion. … Families were not informed in advance that the intent of the observer was to examine television-related behavior. They were told that the observer was interested in studying 'family life' for a college class" (p. 804).

The study of Papper et al. (2004) showed that it is not easy to find families willing to participate in a rather intrusive observation study with unknown observers, even when they are paid. Eventually 101 families participated (out of 14,321 telephone calls). "Observers accompanied participants to work and other venues. They were instructed to minimize casual conversation with participants. Participants were called after observation to complete the observational record for portions of the waking day before observer arrival or after observer departure" (p. 15).

Buijzen and Valkenburg (2008) presented a fine example of an observation study "in the wild." They studied purchase-related communication of parent and child (between newborn and 12) in stores. A total of 269 parent-child dyads were observed by three observers selected after a training session. "The observers followed the parent-child dyad from entering the store to passing through the checkout counter, writing down all behaviours and interactions as they occurred during the store visit. After the parents and children packed their purchases, the observer approached them, informed them about the observation of their store visit, told them about the nature of the study, and asked them for consent to use the observational data. Additionally, the observer asked the parent to fill out a questionnaire" (p. 58). The observations — verbal and nonverbal influence behavior such as demanding, begging, crying, and showing anger — were coded by two coders.

Observation with a hidden camera was employed by Christ and Medoff (1984), a method that was combined with visual recording of content in the study of Brasel and Gips (2011), see Table 6.2.

Table 6.2 Observation protocol and data analysis in the study by Brasel & Gips (2011)

"The behavior of the participants was recorded at 30 frames per second with two unobtrusive video cameras. One of these cameras was focused on the head and eyes of the participants; because the television was located in a raised position (roughly 5 feet off the ground) relative to the laptop screen (which was at desk level), head and eye movement revealed the locus of participants' attention between the two screens. The second video camera was located behind and to the side of the participant to record the television and Internet content chosen. After the 30 minutes of media usage, the television and laptop were shut down, and the participants completed a postsurvey on the experience. Research assistants transformed the raw videos of participant behavior into data files suitable for analysis. Each frame from the video was coded as to whether the participant was looking at the television, the computer, or (rarely) somewhere else. Switches between these states were also coded. From these, participants' gaze durations were computed. Opening new Web pages on the computer and changing channels on the television were also noted. Although stimulus exposure lasted 30 minutes, video records were truncated at 27.5 minutes to eliminate changes in behavior that might result from the anticipated end of the stimulus presentation."

Source: Brasel and Gips (2011, p. 529).

The observation of behavior online including exposure to media is also at the heart of digital ethnography methods. In this strand of research, researchers aim to make sense of digital life through observing and interpreting everyday realities and lived experiences (Pink et al., 2016). These observations are often conducted through participating and interacting with online communities, for example through virtual and augmented reality ethnography, in which researchers immerse themselves in digital environments (e.g., Harrington, 2020), or netnography, a method specifically designed for studying online communities and cultures (see Bartl et al., 2016 for an overview). For example, Achmad et al. (2020) studied the maintenance of local identities and culture through radio stations that focused on campasuri music in East Java. In their study they combined participant observations and interviews with an analysis of digital data such as social media content, conversations on digital personal messaging services, and online discussion boards.

Frequently, digital ethnography relies on the study of digital trace data, as presented in Chapter 5. However, when large-scale digital datasets are analyzed to identify patterns and trends, they are often complemented by qualitative data from participant observation, interviews, and other traditional ethnographic methods (Brooker, 2022). Digital ethnography also includes methods of self-observation. For example, Schellewald (2021) conducted a digital ethnography on the platform TikTok by engaging with the landing page for thirty to sixty minutes a day for six months, emulating typical user behavior, while taking detailed notes in a field diary.

While digital methods alleviate some of the challenges of offline observations, such as the obtrusiveness of the method, they also create new challenges. First, they create unique ethical concerns, in particular related to informed consent, privacy, and data protection. Second and connected to the first challenge are issues pertaining to access to online communities of interest, particularly in cases where privacy settings or community guidelines restrict outsider participation. Third, the multimodality of online spaces that consist of texts, images, and videos, as well as the vast amount of data available, challenges researchers' ability to safely store, and develop, strategies to analyze the data. Fourth and final, digital spaces are becoming increasingly dynamic and personalized. Therefore, when observing online spaces, researchers need to develop strategies to account for the fact that different users are engaging with different interfaces while interacting with each other, implying that what researchers experience on their screen is not the same as what participants experience, who all also are engaging with different experiences.

6.2 Concluding remarks

The above examples illustrate the pros and cons of observation as a method to measure media exposure. On the plus side, the method can give an accurate and detailed picture of real behavior in a naturalistic environment. On the minus side: despite the natural setting, the presence of an observer is unnatural, it is difficult to obtain a representative sample of respondents and situations, studies are time consuming and (therefore) costly, there are ethical concerns, and the comparability, validity and reliability of the coding is questionable.

References

Achmad, Z. A., Ida, R., & Mustain, M. (2020). A virtual ethnography study: The role of cultural radios in Campursari music proliferation in East Java. *ETNOSIA: Journal Etnografi Indonesia*, *5*(2), 221–237. https://doi.org/10.31947/etnosia.v5i2.9787

Araujo, T., & Neijens, P. (2020). Unobtrusive measures for media research. In J. van den Bulck (Ed.), *The international encyclopedia of media psychology* (pp. 1872–1878). Wiley. https://doi.org/10.1002/9781119011071.iemp0049

Bartl, M., Kannan, V. K., & Stockinger, H. (2016). A review and analysis of literature on netnography research. *International Journal of Technology Marketing*, *11*(2), 165–196. https://doi.org/10.1504/IJTMKT.2016.075687

Brasel S. A, & Gips J. (2011). Media multitasking behavior: Concurrent television and computer usage. *Cyberpsychology, Behavior, and Social Networking*, *14*(9), 527–534. https://doi.org/10.1089/cyber.2010.0350

Brooker, P. (2022). Computational ethnography: A view from sociology. *Big Data & Society*, *9*(1). https://doi.org/10.1177/20539517211069892

Buijzen, M., & Valkenburg, P. M. (2008). Observing purchase-related parent-child communication in retail environments: A developmental and socialization perspective. *Human Communication Research*, *34*(1), 50–69. https://doi.org/10.1111/j.1468-2958.2007.00313.x

Christ, W. G., & Medoff, N. J. (1984). Affective state and the selective exposure to and use of television. *Journal of Broadcasting*, *28*(1), 51–63. https://doi.org/10.1080/08838158409386514

Harrington, M. C. (2020). Observation of presence in an ecologically valid ethnographic study using an immersive augmented reality virtual diorama application. *Proceedings of the 2020 IEEE Conference on Virtual Reality and 3D User Interfaces Abstracts and Workshops* (*VRW*), *USA*, 812–813. https://doi.org/10.1109/VRW50115.2020.00257

Krugman, D. M., Cameron, G. T., & White, C. M. (1995). Visual attention to programming and commercials: The use of in-home observations. *Journal of Advertising*, *24*(1), 1–12. https://doi.org/10.1080/00913367.1995.10673464

Lull, J. (1982). How families select television programs: A mass observational study. *Journal of Broadcasting*, *26*(4), 801–811. https://doi.org/10.1080/08838158209364049

Lull, J. (1988). *World families watch television*. Sage.

Papper, R. A, Holmes, M. E., & Popovich, M. N. (2004). Middletown media studies: Media multitasking and how much people really use the media. *International Digital Media and Arts Association Journal*, *1*, 9–50.

Pink, S., Horst, H. A., Postill, J., Hjorth, L., Lewis, T., & Tacchi, J. (2016). *Digital ethnography: Principles and practice*. Sage.

Schellewald, A. (2021). Communicative forms on TikTok: Perspectives from digital ethnography. *International Journal of Communication*, *15*, 1437–1457. https://ijoc.org/index.php/ijoc/article/view/16414

Steiner, G. A. (1966). The people look at commercials: A study of audience behavior. *The Journal of Business*, *39*(2), 272–304. https://doi.org/10.1086/294852

About the authors

Peter Neijens is an Honorary Fellow and Emeritus Professor at ASCoR. He held the Chair in Media and Persuasion at the University of Amsterdam until 2019. His research focuses on persuasive communication, in particular media use and effects in advertising and public opinion.

Theo Araujo is a Full Professor of Media, Organisations and Society, and Scientific Director of ASCoR. His research investigates the dynamic interplay between media and organizations, and what it means for society, with a special focus on trust and technology. His research has also a methodological focus on computational communication science and the implementation of large-scale data collection and analysis for communication research.

Judith Möller holds the chair of "Empirical Communication Research, especially Media Use and Social Media Effects" at the University of Hamburg and the Leibniz Institute for Media Research | Hans Bredow Institute (HBI). Before 2023 she was an Associate Professor of Political Communication at the Amsterdam School of Communication Research.

Claes H. de Vreese is a Distinguished University Professor of AI and Society with a special emphasis on media and democracy. He co-directs the Gravitation Program AlgoSoc and the AI, Media and Democracy Lab. He holds the chair in Political Communication at the Amsterdam School of Communication Research.

7. Eye Tracking

Claire M. Segijn and Emily Vraga

Abstract
In this chapter we discuss eye-tracking measures, terms and metrics, hardware and measurement (in section 7.2), their pros and cons (section 7.3), and quality (section 7.4). Concluding remarks can be found in section 7.5.

Keywords: Area of Interest (AOI), exposure, attention, quality issues, triangulation, reporting

7.1 Introduction

Eye tracking is a technique that is used to measure exposure and attention to stimuli (i.e., materials used in an eye-tracking study that the participant is exposed to, such as a website, social media page, video, or image). For example, eye tracking has been used in communication science to measure attention to social media websites (Ohme et al., 2021; Vraga et al., 2016a), product review pages (Maslowska et al., 2021), or product placement in videos (Boerman et al., 2015). Eye-tracking research can provide insights into how attention is allocated to different stimuli, as well as different elements within a given stimulus (i.e., Area of Interest, AOI), or the order in which different parts of stimuli are viewed (see King et al., 2019) for an overview of eye-tracking research in Communication Science). A stimulus may or may not contain multiple AOIs. An AOI could be the stimulus as a whole, for example if researchers are interested in how much attention participants pay to a social media feed, but also smaller parts of the stimulus, such as a specific social media post or elements of a post (e.g., likes, header, comments). Additionally, eye tracking could also offer an alternative way to measure media exposure, specifically, by considering two measures: stimuli time and AOI time.

Segijn, C. M., & Vraga, E. (2024). Eye tracking. In P. Neijens, T. Araujo, J. Möller, & C. H. de Vreese, *Measuring exposure and attention to media and communication. Solutions to wicked problems* (pp. 113-127). Amsterdam University Press. https://doi.org/10.5117/9789463723176_ch07

7.2 Types of measures

In this section we discuss eye-tracking concepts and measurement.

Eye-tracking terms and metrics

Stimuli time is the total duration that a participant was prese2,95 nted, and thus exposed, to the stimuli. For example, when a participant is asked to review an image of a social media post, the stimuli time indicates the duration that a participant was exposed to that image of the social media post on the screen. When the participant is in control over the duration of the stimulus, media exposure to static stimuli could differ between participants (i.e., different people spend different amounts of time looking at the same stimulus), or within the participant (i.e., the same person spends different amounts of time looking at different stimuli). For example, in some studies participants can decide how long to look at each stimulus (e.g., review page), and they can press a key on the keyboard to move on to the next stimulus (e.g., next review page; see for example Maslowska et al., 2020).

Additionally, *AOI time* (i.e., total duration that the AOI was present to the participant) could be a measure for media exposure with both static stimuli (i.e., does not move, such as an image) and dynamic stimuli (i.e., stimuli that move, such as a video or scrollable website). For static stimuli, AOI time should be the same as stimulus time. For example, by reviewing the static image of the social media post, data is collected on how long someone was exposed to the post in general, but also how long the participant was exposed to different parts of the social media posts (AOIs), such as the header, text, picture, and number of likes. Note that this is different from how much *attention* a participant allocated to the different AOIs.

With dynamic stimuli, however, the AOI does not have to be visible to the participant for the full duration that they are exposed to the stimulus. This could be the case, for example, when researchers are interested in a specific scene or element of a video, a specific post on a social media feed that only becomes visible after scrolling down, or a (part of a) website page to become visible after clicking through links. The duration of the AOI provides a measure of how long the participant was exposed to the AOI, separate from the stimulus overall.

These measures (i.e., stimuli and AOI time) will be different from total dwell/fixation time as it accounts for when the stimulus is present, not when it garners attention. These measures will be of particular interest

Table 7.1 Overview of eye-tracking terms and metrics

Eye-tracking terms and metrics	Alternative name(s)	Meaning
Terms		
Area of Interest (AOI)		A specific area of the stimulus material of which you want to measure attention or exposure.
Gaze		What the eyes are looking at.
Fixation		A series of gaze points close in time and space (i.e., gaze cluster). Gazes of > 60–80 ms to one location are marked as a fixation depending on the software used or researcher decision (see King et al., 2019 for a discussion).
Saccade		Movement between fixations.
Hertz		Sampling frequency.
Metrics		
Fixation count	*Fixation frequency; number of fixations*	Number of fixations on an AOI.
Fixation duration		Average duration of all fixations detected within an AOI.
Total dwell time		Total sum of all fixation (gaze) durations within an AOI.[1]
Time to first fixation (TTFF)	*Entry time*	Time between the first moment the stimulus is presented and the first time a participant fixates on an AOI.
Return to AOI		Number of times that a participant revisits to the same AOI.
Scanpaths	*Gaze path*	Sequence/pattern of eye-movements across the stimulus.
Saccade length	*Saccade amplitude; saccade distance; saccadic velocity*	"Distance between fixations" (p. 152).
Pupil size	*Pupil dilation; pupil diameter*	"(Change in) the (mean) size of participant pupil(s) diameter(s)" (p. 152).
Blink rate	*Number of eye blinks*	"Frequency of blinking as a function of length of exposure to a stimulus/time on a task" (p. 152).
Stimuli time	*Exposure time to stimuli*	Total duration that a participant was exposed to the stimuli.
AOI time	*Exposure time to AOI*	Total duration that the AOI was present to the participant.

Note: This table is an adaption of the Table 1 by King et al. (2019). [1] Depending on the software, metrics may be measured differently. For example, some software includes saccades in dwell time, while others do not. Also, some software provides the option to retrieve dwell time for gazes or fixations. Therefore, it is important to check the (default) settings of the software used and clearly operationalize the metric.

when stimuli include multiple AOIs, so that people will not be attending to all content even when exposed to it.

Table 7.1 provides an overview of eye-tracking terms and metrics. This table is based on King et al. (2019). We build on this table by adding common eye-tracking terms and additional eye-tracking metrics. Most of the metrics are measurements of attention or attention related variables. We added two measures that can serve as a measure of media exposure depending on whether the stimulus is static or dynamic, which we have already discussed above.

The table provides an overview of what the measures capture. Besides that, researchers have assigned meaning to such measures. Longer dwell time or fixation duration, for example, is seen as an indication of the depth of processing (Rayner, 1998) and pupil size as an indication of affective processing (Partala & Surakka, 2003). However, triangulation is needed to get an understanding of the validity of the eye-tracking measures because the meaning of the metrics may differ depending on the participant, context of the study or stimuli (see section 7.3).

Eye-tracking hardware and measurement

Broadly speaking, there are two types of eye trackers, fixed (screen-based) eye trackers or mobile (glasses) eye trackers. The most used eye trackers are fixed (screen-based) eye trackers. These eye trackers typically consist of a bar attached to a screen and they can track where on the screen someone is looking at. This eye tracker could capture exposure or attention to media content presented on a computer screen, but also other screens such as tablets or smartphones. Alternatively, researchers can make use of a mobile eye tracker (glasses). This eye-tracking device is placed on the participant's head — similar to glasses — and the eyes from the point of view of the participant are tracked. This allows tracking of non-screen-based attention, such as attention to newspapers (Kruikemeier et al., 2018), attention between multiple devices (Segijn et al., 2017), or interpersonal communication (Jongerius et al., 2022). Additionally, a development is using people's own devices with cameras for eye tracking purposes (e.g., Schröter et al., 2021). This will allow for online studies using eye-tracking methodology. Like any method, using eye tracking requires specialist knowledge (e.g., Bulling et al., 2020; Duchowski, 2007; Holmqvist, 2011; Holmqvist et al., 2012).

The type of eye tracker used may have different implications for media exposure and attention measurement. The mobile eye tracker has opportunities to measure media exposure and attention given that it shows data from the point of view of the participant. It not only shows what the participant looked at (i.e., attention) but also what was in the environment and thus what the participant was exposed to. However, measurement of media exposure is limited to what is shown in the world view data that is captured with the front camera of the glasses. An additional recording might be needed to capture more of the environment that the participant was in and to get a better understanding of media exposure.

The mobile eye tracker allows for movement and is not limited to what is shown on a screen. For example, a participant can be asked to walk through an environment (e.g., grocery store; Dalton et al., 2015; Pentus et al., 2020) and afterwards exposure to advertisement or certain products could be analyzed.

Limitations of the mobile eye tracker is that it is less precise compared to a fixed eye tracker (Segijn et al., 2023). This might be of especial concern when interested in measuring media attention to smaller AOIs. For example, a mobile eye tracker would be able to capture whether someone looks at the TV in a living room setting, but a specific element of the TV content is more difficult to capture in a reliable way. Therefore, fixed eye trackers are preferred for questions that involve screens-based contents or attention to details in content. Additionally, with the mobile eye tracker, data between participants depend on the head movements by the participant and therefore the output is unique for each participant. This makes analyses of this type of eye-tracking data more challenging compared to data from a fixed eye tracker (for an overview see Segijn et al., 2023).

7.3 Pros and cons

Eye tracking — like all other methods — offers unique benefits and challenges for measuring media exposure and attention. A major benefit of eye tracking is that it allows researchers to more precisely distinguish between exposure and attention, as articulated in Chapter 2. Specifically, by tracking stimuli time or AOI time versus total dwell time or fixation count (see Table 1), we can disentangle the differences between exposure (i.e., the amount of time a stimulus or AOI is on the screen) and attention (i.e., the amount of time or number of fixations that an individual attends to an AOI) for time stamps with recorded fixations, as well as consider their distinct effects on outcomes when appropriate.

A challenge in measuring media exposure, however, is that an (eye-tracking) experiment often involves forced exposure. In most eye tracking studies, participants are asked to review a particular, predetermined set of stimuli and therefore exposure time might be less informative than when measured in an observational (eye-tracking) study.

Similarly, eye tracking typically occurs in a laboratory setting, with specialized equipment such as fixed computer monitors, which may not resemble how people increasingly consume media content. Even those eye-tracking studies that use mobile devices like glasses require participants to consciously put on specific equipment, reinforcing the knowledge that their behavior is being monitored. These participants may thus be more motivated to carefully review stimuli compared to a real-life situation, especially for socially desirable stimuli like news content (Kruikemeier et al., 2018). This might have consequences for measuring (incidental) media exposure and attention, by reducing the external validity of the results, although the degree to which these social desirability concerns may be more or less problematic depending on the research question.

But despite these pressures, eye-tracking data can be less sensitive to demand characteristics (cues revealing the research objectives) and social desirability biases that plague self-reported media exposure, as articulated in Chapter 4. It is difficult for participants to control their gaze, especially for highly salient stimuli, so-called "bottom up" processing (e.g., Buschman & Miller, 2007; Mancas, 2009; Pieters & Wedel, 2004, 2007; Vraga et al., 2016a; see also Chapter 2), meaning that people have difficulties controlling their attention, even when aware of monitoring. As such, people may fall into more habitual patterns of visual attention that reflect their interests, rather than remaining concerned about the demand characteristics of the study. This habitual nature of attention means scholars interested in using eye tracking to distinguish between attention and exposure may feel more confident in the validity of their results.

7.4 Quality

Researchers can take several steps to increase the reliability and validity of eye-tracking research to study media exposure or attention. Different elements are more or less important depending on the research questions. For example, the quality of eye-tracking data can be threatened by factors such as the quality and precision of the eye tracker equipment (camera, illumination) and software, the calibration procedures, the experience and skills of the operator, and the participants (who have different eye physiologies, different abilities to follow instructions, glasses, contact lenses, mascara, and eyelashes), and the recording environment (e.g., lighting, distractions, such as noises and movements) (see Holmqvist et al., 2012). We elaborate on a few examples below to give an idea of what to think about when designing an eye-tracking study.

Experimental design

Researchers may choose to design eye-tracking studies that maximize the external validity of the experience. A commonly used technique is simply telling participants that they should review stimuli as they would normally (e.g., Vraga et al., 2016a). Likewise, those interested in attention to a particular AOI may embed it within a longer, more complicated information environment, to limit social desirability pressures. For example, in studies that track attention to a specific social media post that is embedded in a social media feed (e.g., Kim et al., 2021; Ohme et al., 2021). Finally, some studies choose to tell participants that the aim of the study is different from the actual purpose of the study (cover story) with appropriate debriefing procedure at the end of the study. For example, in the study by Maslowska et al. (2021), the researchers were interested in attention to a sponsored social media post. However, the participants were told that the purpose of the study was to test a Facebook feed developed for academic research, and no mention was made about the advertisement of interest. These steps can reduce the threats to external validity in eye-tracking studies, but these threats cannot be fully ameliorated.

Control over exposure time

Some eye-tracking studies decide to control the amount of time that a participant is exposed to stimuli (e.g., Boerman et al., 2015) while other studies let the participant decide how much time they are exposed to stimuli before moving on (e.g., Maslowska et al., 2020). Yet other studies take a third option: asking people to spend a specific amount of time on a page or stimulus, but not requiring it (Kim et al., 2021). This has implications for media exposure measurement. For example, when the exposure time is controlled by the setup of the study with static stimuli (e.g., each participant has two minutes to read a text), then the media exposure should be the same for all participants (unless they looked away from the screen). However, when the participant is in control of how long to look at the stimuli, the exposure time will likely differ per participant.

Stimuli modality

The modality of the stimuli could have implications for media exposure and how that is measured with eye tracking. For example, people could still be exposed to auditory stimuli even when people are looking somewhere else. This is, for example, the case when people are multitasking with two screens and are focusing their visual attention to one screen (e.g., smartphone, tablet), while audiovisual information is displayed on another screen (e.g., TV). Even though the eyes are not fixated on the audiovisual information, the participants can still be exposed to the content through the auditory information (see for example Segijn et al., 2017). Eye-tracking technology is limited to the visual modality, but by combining it with other measures (e.g., self-report measures), it may be possible to gain insights in exposure to information presented in other modalities. It also offers additional opportunities to consider the effects of exposure as separate from those of attention.

AOI characteristics

If a researcher is interested in measuring media exposure in a dynamic media environment (e.g., video, scrollable website), the location of the AOI is something to take into consideration. Researchers may want to think about placing the AOI above vs. below the fold, or whether participants need to click-through or scroll to get to the AOI. An AOI that is presented above the fold might be less informative when interested in media exposure because all participants are exposed to it at the start of the study. However, this could be realistic if you are interested in newspapers and how much exposure or attention articles get depending on their location.

Additionally, when comparing media exposure to different AOIs the amount of effort that is required by the participant to be exposed should be considered. For example, how much effort is required from the participant to reach the AOIs and is the amount of effort (e.g., number of clicks, amount of scrolling) comparable between AOIs?

The size of the AOI can help determine how long it remains on the screen (i.e., exposure), as well as the amount of attention allocated to that stimulus. Larger AOIs tend to produce higher numbers of fixations that fall within this AOI. Therefore, to compare exposure or attention to different AOIs, it is important to have comparative sizes or account for the size in analyses (see for example Maslowska et al., 2021), unless size is a variable of interest.

Additionally, given measurement error, the distance between different elements in the stimuli or the different AOIs could impact the validity of the results when interested in measuring attention. For example, when a participant allocates attention to an AOI, it is possible that the fixation shows up slightly outside the AOI area because of measurement error. This has implications for stimulus design and the analyses of AOI. Regarding measurement, researchers may need to draw AOIs on the stimulus material that are slightly bigger than the AOI they are interested in. However, when multiple AOIs are located closely together, it is possible that a fixation to one AOI means attention to another. Therefore, researchers may want to space out AOIs and have some (white) space around the AOIs. However, this is a trade off with ecological validity when real-life stimuli (e.g., news site, social media post) do not have that much white space naturally.

Finally, AOI content (e.g., aesthetics, dynamic/static, visual/text) could have implications for media exposure and attention measurements. Different features may attract more attention (e.g., dynamic content, bright colors, vivid or novel content, human faces) than others, and may impact certain eye-tracking metrics, such as time to first fixation (TTFF). Additionally, differences in content could explain fixation time because it takes more/less time or cognitive resources to process the information. For example, pictures or social cues (e.g., likes on a social media post) may take less time and cognitive resources to process than textual information. Similarly, some of this information may be able to be processed when it is in the periphery while other information needs to be fixated on in order to process. This has implications for conclusions that can be drawn.

Hertz (sampling frequency)

Hertz is the sampling interval of the eye-tracking data. For example, 60 Hertz means a sampling interval of 17.67 milliseconds and 120 hertz means a sampling interval of 8.33 milliseconds (Tobii, 2023). Thus, the higher the sampling frequency, the higher the precision of the eye-tracking data (King et al., 2019). On the flipside, higher hertz requires more expensive equipment, better lighting to capture the data, larger data files, and potentially more noise captures (Tobii, 2023). Most eye-tracking studies in Communication Science report using 60–120 Hz (King et al., 2019), although the majority of the eye-tracking studies do not report their sampling rate.

Reporting

This section provided an overview of factors that could influence data quality indicators. To assess the reliability and validity of eye-tracking data, it is important to be consistent in reporting of eye-tracking studies. However, data quality indicators such as accuracy (the difference between true and recorded gaze direction) and precision (consistency of calculated gaze points when the true gaze direction is constant) (tested with an artificial eye), screen position, sizes and margins of AOI, stimuli presentation duration, fixation threshold, sampling rate, and lost data (when glasses, contact lenses, eyelashes, or blinks prevent the video camera from capturing a clear image of the eye) are rarely reported in publications (Holmqvist et al., 2012; King et al., 2019). Bol et al. (2016) showed that standardization in measurement and reporting of eye-tracking studies is lacking, hindering cumulative insights.

Triangulation using eye tracking

Triangulation is important to get a better understanding of the meaning of eye-tracking metrics. When a participant fixates on an element for a long time, it may mean that they are engaged with or deeply processing the information they are fixated on. However, it may also mean that the participant is daydreaming or thinking about elements that are not in the visual field (Cummins, 2017; Duchowski, 2007). Similarly, shorter fixation duration to an AOI does not necessarily mean that participants are not paying attention or not processing that information. For example, a picture or symbol may not require a long fixation duration to process the information or may even be processed from the periphery, and participants may still think about the information while attending to other information. Additionally, the interpretation of the eye-tracking measure might be context-dependent. For example, pupil size may also be affected by other factors, such as changes in the amount of light. By combining eye-tracking measures with other measures (e.g., self-report), researchers can get a better understanding of the validity of the eye-tracking measures beyond visual attention.

However, eye tracking offers one of the most reliable and validated proxy for exposure and attention to content (Duchowski, 2007; Vraga et al., 2016a). Previous studies have found eye tracking to be a more objective and reliable measure of attention compared to self-report measures of attention (de Vreese & Neijens, 2016; Vraga et al., 2019), or recall for stimuli materials (Bol et al., 2016; Vraga et al., 2016a). In particular, work by Vraga and colleagues (2016) suggest that levels of recall of exposure are quite low and divorced from actual attention processes. In looking at a simulated Facebook feed, they asked participants after their experience to estimate how much content they saw was on different topics (such as news, personal life, or politics) and had different formats (like pictures, links, and statuses). Their work suggests people inaccurately recall their social media experiences immediately after exposure, over-reporting the prevalence of pictures and underreporting their experience of "personal" posts. Moreover, they find there are few differences in mis-reporting exposure based on how much attention people paid to the various types of posts, reinforcing the value that eye tracking offers to distinguishing between exposure and attention. However, Segijn et al. (2017) found a relatively high correlation between self-reported and eye-tracking measures of attention when participants were asked directly after exposure to TV and tablet (multitasking) how much attention they allocated to the TV and the tablet. Based on these results, it seems that participants may be better able to self-report their attention to general media usage than specific elements in the media content. Alternatively, eye tracking can be combined with content-analysis (e.g., Zillich & Kessler, 2019) to offer a more robust representation of both content availability and attention. These offer only a few examples of the ways in which eye tracking can be used to complement existing research strategies.

7.5 Concluding remarks

Eye-tracking technology is commonly used in Communication Science to measure exposure and attention to static or dynamic stimuli. The type of research question determines the type of eye-tracking hardware (e.g., fixed or mobile eye tracker) is most suitable and what choices need to be made in the study and stimulus design. These choices have implications for the reliability and validity of the results. This chapter provided some first insights on these implications and the advantages and disadvantages of using eye-tracking methodology to measure media exposure and attention. Eye tracking captures how the eyes move and what they fixate on. Future research is needed to expand our knowledge on the usage of eye trackers for media exposure and how it relates to other measures (e.g., self-reports, digital trace data).

References

Boerman, S. C., van Reijmersdal, E. A., & Neijens, P. C. (2015). Using eye tracking to understand the effects of brand placement disclosure types in television programs. *Journal of Advertising*, *44*(3), 196–207. https://doi.org/10.1080/00913367.2014.967423

Bol, N., Boerman, S. C., Romano Bergstrom, J., & Kruikemeier, S. (2016). An overview of how eye tracking is used in communication research. In M. Antona & C. Stephanidis (Eds.), *Lecture notes in computer science: Vol. 9737. Universal access in human-computer interaction: Methods, techniques, and best practices* (pp. 421–429). Springer International Publishing. https://doi.org/10.1007/978-3-319-40250-5_40

Bulling, A., Huckauf, A., & Jain, E. (Eds.). (2020). *ETRA'20 short papers: ACM symposium on eye tracking research and applications*. Association for Computing Machinery.

Buschman, T. J., & Miller, E. K. (2007). Top-down versus bottom-up control of attention in the prefrontal and posterior parietal cortices. *Science*, *315*(5820), 1860–1862. https://doi.org/10.1126/science.1138071

Cummins, R.G. (2017). Eye tracking. In J. Matthes, C. S. Davis, & R. F. Potter (Eds.), *The international encyclopedia of communication research methods*. https://doi.org/10.1002/9781118901731.iecrm0099

Dalton, N. S., Collins, E., & Marshall, P. (2015). Display blindness? Looking again at the visibility of situated displays using eye tracking. In B. Begole & J. Kim (Eds.), *Proceedings of the 33rd annual ACM conference on human factors in computing systems* (pp. 3889–3898). Association for Computing Machinery. https://doi.org/10.1145/2702123.2702150

de Vreese, C. H., & Neijens, P. (2016). Measuring media exposure in a changing communications environment. *Communication Methods and Measures*, *10*(2–3), 69–80. https://doi.org/10.1080/19312458.2016.1150441

Duchowski, A. (2007). *Eye tracking methodology: Theory and practice* (3rd ed.). Springer International Publishing.

Holmqvist, K. (2011). *Eye tracking: A comprehensive guide to methods and measures*. Oxford University Press.

Holmqvist, K., Nyström, M., & Mulvey, F. (2012). Eye tracker data quality: What it is and how to measure it. In S. N. Spencer (Ed.), *Proceedings of the symposium on eye tracking research and applications* (*ETRA'12*) (pp. 42–52). Association for Computing Machinery. https://doi.org/10.1145/2168556.2168563

Jongerius, C., Twisk, J. W., Romijn, J. A., Callemein, T., Goedemé, T., Smets, E., & Hillen, M. A. (2022). The influence of face gaze by physicians on patient trust:

An observational study. *Journal of General Internal Medicine*, *37*(6), 1408–1414. https://doi.org/10.1007/s11606-021-06906-2

Kim, S. C., Vraga, E. K., & Cook, J. (2021). An eye tracking approach to understanding misinformation and correction strategies on social media: The mediating role of attention and credibility to reduce HPV vaccine misperceptions. *Health Communication*, *36*(13), 1687–1696. https://doi.org/10.1080/10410236.2020.1787933

King, A. J., Bol, N., Cummins, R. G., & John, K. K. (2019). Improving visual behavior research in Communication Science: An overview, review, and reporting recommendations for using eye-tracking methods. *Communication Methods & Measures*, *13*(3), 149–177. https://doi.org/10.1080/19312458.2018.1558194

Kruikemeier, S., Lecheler, S., & Boyer, M. M. (2018). Learning from news on different media platforms: An eye-tracking experiment. *Political Communication*, *35*(1), 75–96. https://doi.org/10.1080/10584609.2017.1388310

Mancas, M. (2009). Relative influence of bottom-up and top-down attention. In L. Paletta & J. K. Tsotsos (Eds.), *Lecture notes in computer science: Vol. 5395. Attention in cognitive systems* (pp. 212–226). Springer International Publishing. https://doi.org/10.1007/978-3-642-00582-4_16

Maslowska, E., Ohme, J., & Segijn, C. M. M. (2021). Attention to social media ads: The role of consumer recommendations and smartphones. *Journal of Interactive Advertising*, *21*(3), 283–296. https://doi.org/10.1080/15252019.2021.1997675

Maslowska, E., Segijn, C. M., Vakeel, K. A., & Viswanathan, V. (2020). How consumers attend to online reviews: An eye-tracking and network analysis approach. *International Journal of Advertising*, *39*(2), 282–306. https://doi.org/10.1080/02650487.2019.1617651

Ohme, J., Araujo, T., de Vreese, C. H., & Piotrowski, J. T. (2021). Mobile data donations: Assessing self-report accuracy and sample biases with the iOS Screen Time function. *Mobile Media & Communication*, *9*(2), 293–313. https://doi.org/10.1177/2050157920959106

Partala, T., & Surakka, V. (2003). Pupil size variation as an indication of affective processing. *International Journal of Human-Computer Studies*, *59*(1–2), 185–198. https://doi.org/10.1016/S1071-5819(03)00017-X

Pentus, K., Ploom, K., Mehine, T., Koiv, M., Tempel, A., & Kuusik, A. (2020). Mobile and stationary eye tracking comparison-package design and in-store results. *Journal of Consumer Marketing*, *37*(3), 259–269. https://doi.org/10.1108/JCM-04-2019-3190

Pieters, R. F. G. M., & Wedel, M. (2004). Attention capture and transfer in advertising: Brand, pictorial, and text-size effects. *Journal of Marketing*, *68*(2), 36–50. https://doi.org/10.1509/jmkg.68.2.36.27794

Pieters, R. F. G. M., & Wedel, M. (2007). Goal control of attention to advertising: The Yarbus implication. *Journal of Consumer Research*, *34*(2), 224–233. https://doi.org/10.1086/519150

Rayner, K. (1998). Eye movements in reading and information processing: 20 years of research. *Psychological Bulletin*, *124*(3), 372–422. https://doi.org/10.1037/0033-2909.124.3.372

Schröter, I., Grillo, N. R., Limpak, M. K., Mestiri, B., Osthold, B., Sebti, F., & Mergenthaler, M. (2021). Webcam eye tracking for monitoring visual attention in hypothetical online shopping tasks. *Applied Sciences*, *11*(19), Article 9281. https://doi.org/10.3390/app11199281

Segijn, C. M., Menheer, P., Hofelich Mohr, A., Lee, G., Kim, E., & Olsen, D. (2023, May 25–29). *Automated object detection in mobile eye-tracking research: Comparing manual coding with tag detection, shape detection, and machine learning* [Paper presentation]. International Communication Association Annual Conference, Toronto, Canada.

Segijn, C. M., Voorveld, H. A. M., Vandeberg, L., & Smit, E. G. (2017). The battle of the screens: Unraveling attention allocation and memory effects when multiscreening. *Human Communication Research*, *43*(2), 295–314. https://doi.org/10.1111/hcre.12106

Tobii. (2023, April 24). *Eye tracker sampling frequency*. https://connect.tobii.com/s/article/eye-tracker-sampling-frequency?language=en_US

Vraga, E. K., Bode, L., Smithson, A. B., & Troller-Renfree, S. (2019). Accidentally attentive: Comparing visual, close-ended, and open-ended measures of attention on social media. *Computers in Human Behavior, 99*, 235–244. https://doi.org/10.1016/j.chb.2019.05.017

Vraga, E., Bode, L., & Troller-Renfree, S. (2016). Beyond self-reports: Using eye tracking to measure topic and style differences in attention to social media content. *Communication Methods and Measures*, *10*(2–3), 149–164. https://doi.org/10.1080/19312458.2016.1150443

Zillich, A. F., & Kessler, S. H. (2019). Measuring selective exposure to online information. Combining eye tracking and content analysis of users' actual search behavior. In C. Peter, T. Naab, & R. Kühne (Eds.), *Methoden und forschungslogik der kommunikationswissenschaft: Vol. 14. Measuring media use and exposure: Recent developments and challenges* (pp. 196–220). Herbert von Halem Verlag.

About the authors

Claire M. Segijn is an Associate Professor and Mithun Program Fellow in Advertising at the University of Minnesota. Her research focuses on the intended and unintended effects of personalized communication and other forms of online surveillance. She uses different quantitative and qualitative methods with a special interest in eye-tracking methodology.

Emily Vraga is an Associate Professor at the University of Minnesota's Hubbard School of Journalism and Mass Communication. Her research tests methods to identify and correct misinformation on social media, to build news literacy to improve audience resiliency, and to encourage attention to more diverse content online.

8. Neurobiological Measures

Frederic R. Hopp and Bert N. Bakker

Abstract
We discuss what neurobiological measures can (and cannot) reveal about media exposure (section 8.2), followed by a discussion of recurrent validity and reliability issues and attempts towards their mitigation (section 8.3). In the end, we briefly discuss how technological advancements and rising societal questions may shape the future of psychophysiological media research.

Keywords: EEG, fMRI, fEMG, heart rate, skin conductance, validity

8.1 Introduction

Media exposure is an inherently dynamic process (A. Lang, 2000; A. Lang & Ewoldsen, 2010). Whether we are watching TV, listening to music, or scrolling through the latest news feed, our brains continuously parse incoming sensory information into coherent, interconnected, and meaningful units. Notably, these complex computational transformations often appear intuitively and without conscious awareness, owing to the evolved, functionally-specialized circuits that constitute our neurobiological make-up (Tooby & Cosmides, 2001).

The conceptualization of media exposure as a dynamic process enabled and orchestrated by an ensemble of biological systems — has sparked an exciting, interdisciplinary field of research (for overviews, see Bolls et al., 2019; Floyd & Weber, 2020). At the same time, the paradigm shift from behaviorism towards information processing (Chaffee, 1980) has also raised an intricate methodological debate: Can we reliably measure the opaque cognitive processing of media content in real time? In this chapter, we review efforts to tackle this question, providing the reader with a critical discussion of psychophysiological measurements that monitor individuals

Hopp, F. R., & Bakker, B. N. (2024). Neurobiological measures. In P. Neijens, T. Araujo, J. Möller, & C. H. de Vreese, *Measuring exposure and attention to media and communication. Solutions to wicked problems* (pp. 129-144). Amsterdam University Press. https://doi.org/10.5117/9789463723176_ch08

during media exposure. Rather than developing a comprehensive introduction and inventory (see R. Potter & Bolls, 2012), we herein take a broader perspective, focusing on what dynamic measures can (and cannot) reveal about media exposure.

8.2 Types of measures

According to the neurophysiological perspective (NPP, Weber et al., 2008), the brain is the central information processing system that parses and integrates incoming sensory sensations and thus provides a powerful gateway for observing how media content is processed in real time. While some studies directly relate brain activity to media processing (see EEG/fMRI below), other prominent psychophysiological measures follow a more indirect route. Here, perception, thinking, feeling, or consciousness are conceptualized as side-effects of an "embodied brain," and we can measure these side effects of brain function using psychophysiological tools (A. Lang et al., 2009). For example, think of the last time you watched a suspenseful scene from an action movie. Perhaps your hands were starting to become sweaty, and your heart rate may have even decelerated slightly. In turn, psychophysiological researchers are interested in how *physiological* changes (e.g., skin conductance and heart rate) index *psychological* events (e.g., arousal, fear) that occurred during media exposure. Taken together, adopting a neurobiological approach to media exposure implies that body, brain, and media jointly act and interact over time, and that describing the interactions between these complex, dynamic systems can reveal how media content gives rise to distinct psychological responses (A. Lang et al., 2009). Following this perspective, we now provide a selection of neurobiological concepts and methods relevant to media exposure research.

Affective responses to media exposure

Affect is a central concept in theories of media exposure such as the Differential Susceptibility to Media Effects Model (Valkenburg & Peter, 2013b). Here, the affective responses are the immediate physiological responses to media exposure. These physiological responses are largely uncontrollable and occur very fast after exposure (between 50 and 100 milliseconds after exposure to a stimulus), while more cognitive evaluations come later (roughly after 500 milliseconds, Lodge & Taber, 2013). These different temporal receptive windows are embedded within a larger processing hierarchy (Hasson et al., 2008) that enables the brain to parse and integrate incoming sensory information (e.g., sounds to words, words to phrases, phrases to paragraphs, etc.).

The circumplex model of affect (Russell, 1980) distinguishes arousal and valence as the two core affective dimensions. Arousal captures the intensity of affect and can be indexed with skin conductance. Valence captures the direction (positive or negative) of affect and may be measured with facial Electromyography (fEMG, see below). Here, we discuss the core characteristics and applications in media exposure research for two affective dimensions: arousal and valence.

Arousal captured with skin conductance. Skin conductance captures the activity of the sympathetic nervous system which is part of the autonomic nervous system (Dawson et al., 2017). Skin conductance measures the varying electrical properties of the skin in response to the increase of sweat secretion in the eccrine glands (for an introduction, see Soroka, 2019). With more sweat secretion, the conductance of electricity improves, and skin conductance levels rise. This increase in skin conductance is interpreted as an increase in arousal. Skin conductance has been used to capture heightened arousal in response to negative images (P. Lang et al., 1993), negative news (Soroka et al., 2019; Soroka & McAdams, 2015), uncivil political debates (Mutz, 2007; Mutz & Reeves, 2005), tobacco ads (Leshner et al., 2022; Clayton, 2022), self-transcendent videos (Clayton et al., 2021), group eating food cues (Liu & Bailey, 2021), political rhetoric (Boyer, 2021), and negative political ads (Wang et al., 2014). Soroka and colleagues, for instance, showed that negative vs. positive news caused an increase in physiological arousal (Soroka et al., 2019; Soroka & McAdams, 2015). Noteworthy, the results of Soroka et al. (2019) were largely uniform for samples of 17 countries across all six continents. In another study, Bakker, Schumacher, et al. (2021) found that listening and watching short messages about contemporary societal issues like immigration, climate change, and inequality increases physiological arousal and that this is especially the case among people with more extreme political attitudes. To summarize, skin conductance is a useful tool to capture arousal during media exposure.

Valence captured with facial Electromyography. The second dimension of affect in the circumplex model is valence. Valence can be captured using facial Electromyography (fEMG). Specifically, tiny (not observable) muscle contractions in a specific facial region induce bioelectric signals at the surface of the skin in the face. fEMG captures these bioelectric signals (R. Potter & Bolls, 2012). An advantage of fEMG is that it detects rapid changes in the contractions of facial muscles over time (van Boxtel, 2010), making it possible to link media processing and physiological responses at a high temporal resolution.

The *corrugator supercilii* muscle region (above the eyebrow) captures negative affect. Corrugator activity has been recorded in response to negative images (P. Lang et al., 1993), and negative words (Hietanen et al., 1998). The *zygomaticus major* muscle region (pulls the corners of the mouth into a smile, Larsen et al., 2003), captures positive affect. Zygomaticus activity increases in response to positively valenced images (van Oyen Witvliet & Vrana, 1995) and videos (Cacioppo et al.,1986). Other muscles have been associated with specific discrete emotions, such as the *levator labii*, which is associated with disgust (Chapman et al., 2009). Bakker, Schumacher, and Homan (2020), for instance, showed that participants had strong labii responses to a politician from the out-party versus the in-party. Moreover, participants experienced an increase in labii activity once a politician of the in-party (compared to the out-party) committed a moral violation.

Along these lines, studies exposing people to various forms of media content show that negative affect in media content causes corrugator activity. For instance, listing to radio advertisements, listeners showed a greater corrugator response "following the onset of negatively valenced words compared to positively valenced words" (Lee & R. Potter, 2020, p. 1154). Likewise, watching highway safety videos featuring a high aversive tone increased corrugator activity compared to a video with a low aversive tone (Howell et al., 2018).

Yet, when the corrugator activity is paired to what people say they experience, a different pattern emerges. Homan et al. (2022), for instance, find that participants relaxed their corrugator muscle in response to in-party politicians' emotional displays, irrespective of the emotion the in-party politician displayed. But participants report feeling significantly angrier when watching an in-party politician display anger compared to a neutral or happy expression of the same politician. Similarly, Bakker, Schumacher, and Rooduijn (2021) showed that listening to incongruent messages lead to an increase in corrugator activity. This corrugator activity was basically uncorrelated with self-reported discrete emotions, but the increase in corrugatory activity predicated post-treatment attitude change.

To summarize, fEMG seems especially relevant in contexts in which participants have no or suppressed cognitive access to their affective responses, or in cases where they may be unwilling to report their views (e.g., due to social desirability). This may be quite common in media exposure studies, especially when dealing with contested topics like politics in modern Western democracies.

Attention captured with heart rate variability

Attention to specific content features naturally is of core interest to researchers interested in understanding how media users dynamically register incoming information (Fisher et al., 2023). We focus our discussion on heart rate variability here. In media exposure research, heart rate variability is used to measure (cognitive) attention (A. Lang et al., 1996; Soroka et al., 2019; Soroka & McAdams, 2015)[1]. A few recent studies illustrate the potential use of heart rate variability — as a measure of attention — for media exposure. Soroka and colleagues find that there is an increase in heart rate variability in response to negative vs. positive news (Soroka et al., 2016; Soroka et al., 2019; Soroka & McAdams, 2015). The authors interpret these results as evidence that people express more attention to negative vs. positive news. Dunaway and Soroka (2021) asked whether information processing differs when people consume information on a smartphone or on a larger computer screen. Dunaway and Soroka (2021, p. 69) explain that the "results suggest lower levels of cognitive access" as captured with heart rate variability "to video news content on a mobile-sized screen" compared to the normal screen. Dunaway and Soroka (2021) conclude that their study has "potentially important consequences for public attention to current affairs in an increasingly mobile media environment" (p. 69). However, Bakker, Schumacher, and Rooduijn (2021), for instance, found no indication that exposure to counter-attitudinal messages, captured in a campaign-ad style — increases attention (e.g., no statistically significant increase in heart rate variability) — as the theory of motivated reasoning would lead us to expect as counter-arguing takes effort (Lodge & Taber, 2013). Taken together, these studies illustrate heart rate variability is a useful tool for capturing attention to media.

fMRI

In the past decade, functional magnetic resonance imaging (fMRI) has established itself as an appropriate neurophysiological method for examining cognitive responses during dynamic media exposure (for a recent review, see Hopp & Weber, 2020). In a nutshell, fMRI relies on the magnetic properties of hemoglobin to measure the contrast of oxygenated to deoxygenated blood at

1 Sometimes heart variability is also seen as a measure of arousal but any acceleration in heart rate that comes from arousal will be overwhelmed by the deceleration that comes with attention (R. Potter & Bolls, 2012).

a high spatiotemporal resolution, thereby serving as an indirect assessment of neural activity. Before conducting an fMRI study, media researchers typically define a set of brain regions (Regions of Interest [ROIs]) or networks of ROIs whose activity is known to correlate with a psychological concept or process of interest (e.g., aggression, valuation, or attention; but see Poldrack, 2006) and then observe how neural activity in these ROIs dynamically changes in response to certain media content features. Hence, fMRI is a useful tool for continuously interrogating a range of latent cognitive processes relevant to media exposure research, including attention (Huskey et al., 2018), narrative engagement (Grall et al., 2021), valuation (Scholz et al., 2017), counterarguing (Weber et al., 2015), or affective experiences (Chan et al., 2020).

In one of the first studies employing fMRI during media exposure, Weber et al. (2006) showed that playing a violent video game suppressed affective areas of the anterior cingulate cortex (ACC) and the amygdala, highlighting the potential of virtual violence for inducing aggressive cognitions. Almost simultaneously, Hasson et al. (2004) demonstrated that neural activity in the visual cortex is highly correlated across subjects watching the same movie, revealing that individual brains show a surprising tendency to "tick collectively" during natural vision. Since these seminal studies, fMRI has successfully been applied to delineate the neural correlates of counterarguing during anti-drug messages (Weber et al., 2015), the valuation and information virality of news (Scholz et al., 2017), the synchronization of attention and reward brain networks during media induced flow experiences (Huskey et al., 2018), and even the ideologically aligned and idiosyncratic brain responses during politically polarized perception (van Baar et al., 2021).

A key advantage of fMRI over previously described psychophysiological measurements is its *multivariate* recording of neural activity. By examining the neural activity "pattern" (i.e., covariance) within ROIs or even the whole brain, recent work has demonstrated that distinct psychological events, from discrete emotions (Kragel et al., 2019) to moral intuitions (Hopp et al., 2022), can be "decoded" from underlying brain representations. This Multivariate Pattern Analysis (MVPA, Norman et al., 2006) holds great potential to push psychophysiological measurements beyond broader dimensions of cognition and affect (cf. the circumplex model of affect) towards more fine-grained *semantic* interpretations of media content. For example, Hopp (2021) used MVPA and found that Republicans and Democrats in the US experience different moral violations when processing the same political attack advertisements, thereby revealing a "morality bias" that could explain ongoing polarization and out-party animosity.

While the above studies only provide a glimpse of past research, they succinctly demonstrate the method-theory synergy that inspires the exciting field of media neuroscience (Weber et al., 2018). At the same time, access to MRI facilities, average scanning costs, and a steep technical learning curve have hindered communication researchers from incorporating fMRI into their methodological arsenal. Yet, powerful and cost-effective alternatives to fMRI have been applied in media exposure research, including electroencephalography (EEG, Morey, 2018) and functional near infrared spectroscopy (fNIRS, Dieffenbach et al., 2021). We believe that fMRI is a useful tool for our understanding of the effects of media exposure on individuals.

8.3 The perils of physiological measures for media exposure research

In the previous section, we reviewed how psychophysiological measurements, informed by theoretical advancements in biology and cognitive neuroscience, have pushed the envelope of media exposure research. At the same time, psychophysiological tools have their respective limitations and offer no panacea for studying media exposure. Accordingly, we now turn to a discussion of the common perils that have stymied psychophysiological research, and offer several solutions to minimize erroneous interpretations and questionable research practices (QRPs, Bakker, Jaidka, et al., 2021).

Methodological considerations: validity, reliability and open science

Advancements in engineering will naturally continue to improve the fidelity and reliability of psychophysiological measurements. However, we must not forget that the opportunities of neurobiological measures for advancing media exposure research are fundamentally constrained by our understanding (and experimental control) of the independent variable: media *content*. Thus, it is imperative that in designing our studies, we make sure to, by design, rule out potential confounding variables. One possibility is for the researcher to produce the stimuli (see Bakker, Schumacher, & Rooduijn, 2021). Yet, often media researchers would want to use naturalistic stimuli (see Soroka et al., 2019). In that case it is important, to be rigorous in content analyzing stimuli to rule out confounds that may unknowingly modulate psychophysiological responses (Grall & Finn, 2022). Here, recent developments in computer vision and signal processing offer media researchers powerful, automated tools for extracting a range of audiovisual features known to correlate with psychophysiological responses, including amplitude, scale, or frame centrality (Malik et al., 2022). Combined with a carefully controlled, pre-tested, and pre-registered study design, we increase our power to isolate effects and draw solid causal inferences.

Moreover, a largely unexplored question is the extent to which neurobiological measures and more cognitive and reflective measures such as self-reported thoughts and emotions should align. For instance, one critique towards skin conductance is based upon the fact that physiological responses to threatening images did not — or very weakly — correlate with self-reported affective responses to the same images (see for instance, Osmundsen et al., 2022). This lack of concordance is actually not evidence for the lack of validity of one measure over the other. Physiological measures are tapping into other aspects of affect than the self-reports do (Keltner & Gross, 1999). It is therefore not a question of which is "better"; these measures are simply tapping into different aspects of affect. In fact, we think that one of the big puzzles in the field is to outline when and under which conditions physiological and self-reports align or not (see also, Bakker, Schumacher, Gothreau, & Arceneaux, 2020; Bakker, Schumacher, & Rooduijn, 2021).

At the same time, we encourage scholars to critically assess the predictive validity of the neurobiological measures as well as the self-reported measures. Here there might be different options. First, neurological and self-reported responses have independent effects on the outcome of interest (see for instance Bakker, Schumacher, & Rooduijn, 2021; LeDoux & Pine, 2016). Alternatively, one could argue the predictive validity of the neurobiological responses is the strongest when it is aligned with a more cognitive response (for a similar argument, see Bakker, Schumacher, & Homan, 2020). Others might argue that the self-reported effects are the strongest on other self-reported measures, while more neurobiological measures have stronger effects on other implicit measures (Evers et al., 2014). We do not know what is the correct answer, yet. However, we do advocate for a multi-level, pluralistic approach when understanding the brain–behavior–media relationship (see Marr's tri-level framework, Huskey et al., 2020). Behavioral and self-report work is often a necessary first step to understanding *why* a behavior occurs (e.g., selecting negative over positive news), after which psychophysiology can reveal *how* this behavior is physically instantiated (e.g., heightened attention and arousal towards negative information).

Furthermore, it will not come as a surprise to the reader of this chapter (and book) that laboratory-based studies lack ecological validity (Bronfenbrenner, 1979; Nastase et al., 2020). It is not natural to read, watch or listen to news (or other media content) while sitting in an artificial (almost hospital-like) laboratory environment while connected to a set of wires to your body (in the case of physiology) or even laying down in a large and noisy magnet (in case of fMRI). Technological advancements, however, have led to an increase in wearable equipment. As a consequence, neurobiological measures can now be collected in more natural environments: wristbands can for instance capture physiological arousal (Bolinski et al., 2021; Konvalinka et al., 2011). Using these and other wearable devices, it becomes possible to study people at home, at school, among their friends, or at any other place where people process media content. We think there is a lot of unexplored ground to study the neurobiological responses to media exposure outside the classical laboratory setting. If and how the conclusions of studies discussed in this chapter will differ, or will remain the same, is an open-ended question. Regardless of the answer, we think such ecological studies will make important contributions to our understanding of the neurobiological responses to media exposure.

Finally, media exposure studies that rely upon neurobiological measures need replications of seminal, foundational studies to learn more about their replicability and/or generalizability (McEwan et al., 2018). We also encourage scholars to preregister their hypotheses, design, sample size justification, and analysis strategy. A priori power analyses — and more in general sample size justification — should become the standard (Lakens, 2022). Adopting these open science practices will lead to more studies that can reliably detect the (often small) effects of interest. Ultimately, adopting this will increase the credibility of using neurobiological measures for media exposure research.

8.4 Concluding remarks

In this chapter, we have provided the reader with a rough overview of possible neurobiological theories and methods that can be used in media exposure research. We hope the reader treats this chapter as a starting point. For us, the two most important take-home messages are as follows: First, information processing is too complex to solely rely upon behavioral and self-reported measures. Our body is responding in a much more complex way to media exposure. This leaves us with our second take-home point. The neurobiological responses to media exposure call for interdisciplinary partnerships: psychologists and neuroscientists bring their expertise when it comes to the neurobiological mechanisms, while communication researchers bring their rich understanding of media exposure, coupled with methodological expertise in content analysis. Together, these partnerships have a unique contribution to make. As a consequence, the neurobiological responses during media exposure are too complex to leave to one particular discipline but call for exciting interdisciplinary collaborations.

To conclude, we think that neurobiological theories and methods deserve a place in the toolbox of communication researchers: if we want to study the full complexity of responses to media exposure, then neurobiological theories and methods belong in the toolbox. This requires researchers to learn new theories and new methods or join interdisciplinary collaborations.

References

Bakker, B. N., Jaidka, K., Dörr, T., Fasching, N., & Lelkes, Y. (2021). Questionable and open research practices: Attitudes and perceptions among quantitative communication researchers. *Journal of Communication*, *71*(5), 715–738. https://doi.org/10.1093/joc/jqab031

Bakker, B. N., Schumacher, G., Gothreau, C., & Arceneaux, K. (2020). Conservatives and liberals have similar physiological responses to threats. *Nature Human Behaviour*, *4*(6), 613–621. https://doi.org/10.1038/s41562-020-0823-z

Bakker, B. N., Schumacher, G., & Homan, M. D. (2020). Yikes! Are we disgusted by politicians? *Politics and the Life Sciences*, *39*(2), 135–153. https://doi.org/10.1017/pls.2020.16

Bakker, B. N., Schumacher, G., & Rooduijn, M. (2021). Hot politics? Affective responses to political rhetoric. *American Political Science Review*, *115*(1), 150–164. https://doi.org/10.1017/S0003055420000519

Bolinski, F., Etzelmüller, A., de Witte, N. A. J., van Beurden, C., Debard, G., Bonroy, B., Cuijpers, P., Riper, H., & Kleiboer, A. (2021). Physiological and self-reported arousal in virtual reality versus face-to-face emotional activation and cognitive restructuring in university students: A crossover experimental study using wearable monitoring. *Behaviour Research and Therapy*, *142*, Article 103877. https://doi.org/10.1016/j.brat.2021.103877

Bolls, P. D., Weber, R., Lang, A., & Potter, R. F. (2019). Media psychophysiology and neuroscience: Bringing brain science into media processes and effects research. In M. B. Oliver, A. R. Raney, & J. Bryant (Eds.), *Media effects: Advances in theory and research* (pp. 195–210). Routledge.

Boyer, M. M. (2021). Aroused argumentation: How the news exacerbates motivated reasoning. *The International Journal of Press/Politics*, *28*(1), 92–115. https://doi.org/10.1177/19401612211010577

Bronfenbrenner, U. (1979). *The ecology of human development: Experiments by nature and design*. Harvard University Press.

Cacioppo, J., Petty, R. E., Losch, M. E., & Kim, H. S. (1986). Electromyographic activity over facial muscle regions can differentiate the valence and intensity of affective reactions. *Journal of Personality and Social Psychology*, *50*(2), 260–268. https://doi.org/10.1037/0022-3514.50.2.260

Chaffee, S. H. (1980). Mass media effects: New research perspectives. In G. C. Wilhoit & H. de Bock (Eds.), *Mass communication review yearbook, Vol. 1* (pp. 77–108). Sage.

Chan, H. Y., Smidts, A., Schoots, V. C., Sanfey, A. G., & Boksem, M. A. (2020). Decoding dynamic affective responses to naturalistic videos with shared neural patterns. *NeuroImage*, *216*, Article 116618. https://doi.org/10.1016/j.neuroimage.2020.116618

Chapman, H. A., Kim, D. A., Susskind, J. M., & Anderson, A. K. (2009). In bad taste: Evidence for the oral origins of moral disgust. *Science*, *323*(5918), 1222–1226. https://doi.org/10.1126/science.1165565

Clayton, R. B. (2022). On the psychophysiological and defensive nature of psychological reactance theory. *Journal of Communication*, *72*(4), 461–475. https://doi.org/10.1093/joc/jqac016

Clayton, R. B., Raney, A. A., Oliver, M. B., Neumann, D., Janicke-Bowles, S. H., & Dale, K. R. (2021). Feeling transcendent? Measuring psychophysiological responses to self-transcendent media content. *Media Psychology*, *24*(3), 359–384. https://doi.org/10.1080/15213269.2019.1700135

Dawson, M. E., Schell, A. M., & Filion, D. L. (2007). The electrodermal system. In J. T. Cacioppo, L. G. Tassinary, & G. G. Berntson (Eds.), *Handbook of psychophysiology* (pp. 159–181). Cambridge University Press. https://doi.org/10.1017/CBO9780511546396.007

Dieffenbach, M. C., Gillespie, G. S. R., Burns, S. M., McCulloh, I. A., Ames, D. L., Dagher, M. M., Falk, E. B., & Lieberman, M. D. (2021). Neural reference groups: A synchrony-based classification approach for predicting attitudes using fNIRS. *Social Cognitive and Affective Neuroscience*, *16*(1–2), 117–128. https://doi.org/10.1093/scan/nsaa115

Dunaway, J., & Soroka, S. (2021). Smartphone-size screens constrain cognitive access to video news stories. *Information, Communication & Society*, *24*(1), 69–84. https://doi.org/10.1080/1369118X.2019.1631367

Evers, C., Hopp, H., Gross, J. J., Fischer, A. H., Manstead, A. S., & Mauss, I. B. (2014). Emotion response coherence: A dual-process perspective. *Biological Psychology*, *98*, 43–49. https://doi.org/10.1016/j.biopsycho.2013.11.003

Fisher, J. T., Hopp, F. R., & Weber, R. (2023). Mapping attention across multiple media tasks. *Media Psychology*, *26*(5), 505–529. https://doi.org/10.1080/15213269.2022.2161576

Floyd, K., & Weber, R. (2020). *The handbook of communication science and biology*. Routledge.

Grall, C., & Finn, E. S. (2022). Leveraging the power of media to drive cognition: A media-informed approach to naturalistic neuroscience. *Social Cognitive and Affective Neuroscience*, *17*(6), 598–608. https://doi.org/10.1093/scan/nsac019

Grall, C., Tamborini, R., Weber, R., & Schmälzle, R. (2021). Stories collectively engage listeners' brains: Enhanced intersubject correlations during reception of personal narratives. *Journal of Communication*, *71*(2), 332–355. https://doi.org/10.1093/joc/jqab004

Hasson, U., Nir, Y., Levy, I., Fuhrmann, G., & Malach, R. (2004). Intersubject synchronization of cortical activity during natural vision. *Science*, *303*(5664), 1634–1640. https://doi.org/10.1126/science.1089506

Hasson, U., Yang, E., Vallines, I., Heeger, D. J., & Rubin, N. (2008). A hierarchy of temporal receptive windows in human cortex. *Journal of Neuroscience*, *28*(10), 2539–2550. https://doi.org/10.1523/JNEUROSCI.5487-07.2008

Hietanen, J. K., Surakka, V., & Linnankoski, I. (1998). Facial electromyographic responses to vocal affect expressions. *Psychophysiology*, *35*(5), 530–536. https://doi.org/10.1017/S0048577298970445

Homan, D., Maaike, Schumacher, G., & Bakker, N., Bert. (2022). Facing emotional politicians: Do emotional displays of politicians evoke mimicry and emotional contagion?. *Emotion*, *23*(6), 1702–1713. https://doi.org/10.1037/emo0001172

Hopp, F. R. (2021). *Moral idiosynchrony: Variability in naturalistic complexity modulates intersubject representational similarity in moral cognition* (Publication No. 28868128) [Doctoral dissertation, University of California]. ProQuest.

Hopp, F. R., Amir, O., Fisher, J., Grafton, S., Sinnott-Armstrong, W., & Weber, R. (2022). *Moral foundations elicit shared and dissociable cortical activation modulated by political ideology*. Research Square. https://doi.org/10.21203/rs.3.rs-2133317/v1

Hopp, F. R., & Weber, R. (2020). The state of the art and the future of functional magnetic resonance imaging in communication research. In K. Floyd & R. Weber (Eds.), *The handbook of communication science and biology* (pp. 279–292). Routledge.

Howell, M., Ekman, D. S., Almond, A., & Bolls, P. (2018). Switched on: How the timing of aversive content in traffic safety videos impact psychophysiological indicators of message processing. *Health Communication*, *34*(13), 1663–1672. https://doi.org/10.1080/10410236.2018.1517706

Huskey, R., Bue, A. C., Eden, A., Grall, C., Meshi, D., Prena, K., Schmälzle, R., Scholz, C., Turner, B. O., & Wilcox, S. (2020). Marr's tri-level framework integrates biological explanation across communication subfields. *Journal of Communication*, *70*(3), 356–378. https://doi.org/10.1093/joc/jqaa007

Huskey, R., Craighead, B., Miller, M. B., & Weber, R. (2018). Does intrinsic reward motivate cognitive control? A naturalistic-fMRI study based on the synchronization theory of flow. *Cognitive, Affective, & Behavioral Neuroscience*, *18*, 902–924. https://doi.org/10.3758/s13415-018-0612-6

Keltner, D., & Gross, J. J. (1999). Functional accounts of emotions. *Cognition & Emotion*, *13*(5), 467–480. https://doi.org/10.1080/026999399379140

Konvalinka, I., Xygalatas, D., Bulbulia, J., Schjødt, U., Jegindø, E.-M., Wallot, S., van Orden, G., & Roepstorff, A. (2011). Synchronized arousal between performers and related spectators in a fire-walking ritual. *Proceedings of the National Academy of Sciences*, *108*(20), 8514–8519. https://doi.org/10.1073/pnas.1016955108

Kragel, P. A., Reddan, M. C., LaBar, K. S., & Wager, T. D. (2019). Emotion schemas are embedded in the human visual system. *Science Advances*, *5*(7), Article eaaw4358. https://doi.org/10.1126/sciadv.aaw4358

Lakens, D. (2022). Sample size justification. *Collabra: Psychology*, *8*(1), Article 33267. https://doi.org/10.1525/collabra.33267

Lang, A. (2000). The limited capacity model of mediated message processing. *Journal of Communication*, *50*(1), 46–70. https://doi.org/10.1111/j.1460-2466.2000.tb02833.x

Lang, A., & Ewoldsen, D. R. (2010). Beyond effects: Conceptualizing communication as dynamic, complex, nonlinear, and fundamental. In S. Allan (Ed.), *Rethinking communication: Keywords in communication research* (pp. 111–122). Hampton Press.

Lang, A., Newhagen, J., & Reeves, B. (1996). Negative video as structure: Emotion, attention, capacity, and memory. *Journal of Broadcasting & Electronic Media*, *40*(4), 460–477. https://doi.org/10.1080/08838159609364369

Lang, A., Potter, R. F., & Bolls, P. (2009). Where psychophysiology meets the media: Taking the effects out of mass media research. In J. Bryant & M. B. Oliver (Eds.), *Media effects: Advances in theory and research* (pp. 185–206). Routledge.

Lang, P. J., Greenwald, M. K., Bradley, M. M., & Hamm, A. O. (1993). Looking at pictures: Affective, facial, visceral, and behavioral reactions. *Psychophysiology*, *30*(3), 261–273. https://doi.org/10.1111/j.1469-8986.1993.tb03352.x

Larsen, J. T., Norris, C. J., & Cacioppo, J. (2003). Effects of positive and negative affect on electromyographic activity over zygomaticus major and corrugator supercilii. *Psychophysiology*, *40*(5), 776–785. https://doi.org/10.1111/1469-8986.00078

LeDoux, J. E., & Pine, D. S. (2016). Using neuroscience to help understand fear and anxiety: A two-system framework. *American Journal of Psychiatry*, *173*(11), 1083–1093. https://doi.org/10.1176/appi.ajp.2016.16030353

Lee, S., & Potter, R. F. (2020). The impact of emotional words on listeners' emotional and cognitive responses in the context of advertisements. *Communication Research*, *47*(8), 1155–1180. https://doi.org/10.1177/0093650218765523

Leshner, G., Phan, L., Stevens, E. M., Johnson, A. C., Villanti, A. C., Kim, N., Kim, S., Ma, H., Seo, J., Sun, F., Keller-Hamilton, B., Wagener, T. L., & Mays, D. (2021). Psychophysiological responses to hookah tobacco public education messages among young adults. *Health Psychology*, *40*(9), 569–577. https://doi.org/10.1037/hea0001092

Liu, J., & Bailey, R. L. (2021). The influence of social eating norms on motivational responses when processing fast food advertisements. *Journal of Health Communication*, *26*(11), 773–780. https://doi.org/10.1080/10810730.2021.2008550

Lodge, M., & Taber, C. S. (2013). *The rationalizing voter.* Cambridge University Press.

Malik, M., Hopp, F. R., & Weber, R. (2022). Representations of racial minorities in popular movies: A content-analytic synergy of computer vision and network science. *Computational Communication Research*, *4*(1), 208–253. https://doi.org/10.5117/CCR2022.1.006.MALI

McEwan, B., Carpenter, C. J., & Westerman, D. (2018). On replication in communication science. *Communication Studies*, *69*(3), 235–241. https://doi.org/10.1080/10510974.2018.1464938

Morey, A. C. (2018). Electroencephalography in communication research: A review of the past and a glimpse of future possibilities. *Annals of the International Communication Association*, *42*(4), 243–269. https://doi.org/10.1080/23808985.2018.1537723

Mutz, D. C. (2007). Effects of 'in-your-face' television discourse on perceptions of a legitimate opposition. *American Political Science Review*, *101*(4), 621–635. https://doi.org/10.1017/S000305540707044X

Mutz, D. C., & Reeves, B. (2005). The new videomalaise: Effects of televised incivility on political trust. *American Political Science Review*, *99*(1), 1–15. https://doi.org/10.1017/S0003055405051452

Nastase, S. A., Goldstein, A., & Hasson, U. (2020). Keep it real: Rethinking the primacy of experimental control in cognitive neuroscience. *NeuroImage*, *222*, Article 117254. https://doi.org/10.1016/j.neuroimage.2020.117254

Norman, K. A., Polyn, S. M., Detre, G. J., & Haxby, J. V. (2006). Beyond mind-reading: Multi-voxel pattern analysis of fMRI data. *Trends in Cognitive Sciences*, *10*(9), 424–430. https://doi.org/10.1016/j.tics.2006.07.005

Osmundsen, M., Hendry, D. J., Laustsen, L., Smith, K. B., & Petersen, M. B. (2022). The psychophysiology of political ideology: Replications, reanalyses, and recommendations. *The Journal of Politics*, *84*(1), 50–66. https://doi.org/10.1086/714780

Poldrack, R. A. (2006). Can cognitive processes be inferred from neuroimaging data?. *Trends in Cognitive Sciences*, *10*(2), 59–63. https://doi.org/10.1016/j.tics.2005.12.004

Potter, R. F., & Bolls, P. D. (2012). *Psychophysiological measurement and meaning: Cognitive and emotional processing of media.* Routledge.

Russell, J. A. (1980). A circumplex model of affect. *Journal of Personality and Social Psychology*, *39*(6), 1161–1178. https://doi.org/10.1037/h0077714

Scholz, C., Baek, E. C., O'Donnell, M. B., Kim, H. S., Cappella, J. N., & Falk, E. B. (2017). A neural model of valuation and information virality. *Proceedings of the National Academy of Sciences*, *114*(11), 2881–2886. https://doi.org/10.1073/pnas.1615259114

Soroka, S. (2019). Skin conductance in the study of politics and communication. In G. Foster (Ed.), *Biophysical measurement in experimental social science research* (pp. 85–104). Elsevier. https://doi.org/10.1016/B978-0-12-813092-6.00007-1

Soroka, S., Fournier, P., & Nir, L. (2019). Cross-national evidence of a negativity bias in psychophysiological reactions to news. *Proceedings of the National Academy of Sciences*, *116*(38), 18888–18892. https://doi.org/10.1073/pnas.1908369116

Soroka, S., Gidengil, E., Fournier, P., & Nir, L. (2016). Do women and men respond differently to negative news?. *Politics & Gender*, *12*(2), 344–368. https://doi.org/10.1017/S1743923X16000131

Soroka, S., & McAdams, S. (2015). News, politics, and negativity. *Political Communication*, *32*(1), 1–22. https://doi.org/10.1080/10584609.2014.881942

Tooby, J., & Cosmides, L. (2001). Does beauty build adapted minds? Toward an evolutionary theory of aesthetics, fiction, and the arts. *SubStance*, *30*(1–2), 6–27. https://doi.org/10.2307/3685502

Valkenburg, P. M., & Peter, J. (2013b). The differential susceptibility to media effects model. *Journal of Communication*, *63*(2), 221–243. https://doi.org/10.1111/jcom.12024

van Baar, J. M., Halpern, D. J., & FeldmanHall, O. (2021). Intolerance of uncertainty modulates brain-to-brain synchrony during politically polarized perception. *Proceedings of the National Academy of Sciences*, *118*(20), Article e2022491118. https://doi.org/10.1073/pnas.2022491118

van Boxtel, A. (2010). Facial EMG as tool for inferring affective states. In A. Spink, F. Grieco, O. Krips, L. Loijens, L. Noldus, & P. Zimmerman (Eds.), *Proceedings of Measuring Behavior 2010* (pp. 104–108). Noldus Information Technology.

van Oyen Witvliet, C., & Vrana, S. R. (1995). Psychophysiological responses as indices of affective dimensions. *Psychophysiology*, *32*(5), 436–443. https://doi.org/10.1111/j.1469-8986.1995.tb02094.x

Wang, Z., Morey, A. C., & Srivastava, J. (2014). Motivated selective attention during political ad processing: The dynamic interplay between emotional ad content and candidate evaluation. *Communication Research*, *41*(1), 119–156. https://doi.org/10.1177/0093650212441793

Weber, R., Fisher, J. T., Hopp, F. R., & Lonergan, C. (2018). Taking messages into the magnet: Method-theory synergy in communication neuroscience. *Communication Monographs*, *85*(1), 81–102. https://doi.org/10.1080/03637751.2017.1395059

Weber, R., Huskey, R., Mangus, J. M., Westcott-Baker, A., & Turner, B. O. (2015). Neural predictors of message effectiveness during counterarguing in antidrug campaigns. *Communication Monographs*, *82*(1), 4–30. https://doi.org/10.1080/03637751.2014.971414

Weber, R., Ritterfeld, U., & Mathiak, K. (2006). Does playing violent video games induce aggression? Empirical evidence of a functional magnetic resonance imaging study. *Media Psychology*, *8*(1), 39–60. https://doi.org/10.1207/S1532785XMEP0801_4

Weber, R., Sherry, J., & Mathiak, K. (2008). The neurophysiological perspective in mass communication research. Theoretical rationale, methods, and applications. In M. J. Beatty, J. C. McCroskey, & K. Floyd (Eds.), *Biological dimensions of communication: Perspectives, methods, and research* (pp. 41–71). Hampton Press.

About the authors

Bert N. Bakker is an Associate Professor for Good Research Practices at the Amsterdam School of Communication Research (University of Amsterdam). He studies the causes of citizens' political preferences and published in leading journals in communication science, political science, and psychology. Currently, Bert serves as senior associate editor at the Journal of Experimental Political Science.

Frederic R. Hopp is an Assistant Professor at the University of Amsterdam's School of Communication Research. He is interested in the moral content of human communication and how moralized messages are cognitively processed and motivate behavior. His research leverages a method-theory synergy that combines behavioral experiments, neuroimaging, natural language processing, and machine learning.

9. Ecological Measures

Abstract
Ecological measures for MCE "estimate the possibility or opportunity for exposure" (Liu & Hornik, 2016, p. 116). We discuss different types of these measures including their pros and cons in section 9.1. Quality issues are discussed in section 9.2. Conclusions are drawn in section 9.3.

Keywords: media availability, campaign intensity, campaign tracking, field experiments, quality issues

9.1 Types of ecological measures

Ecological measures of media exposure, also known as exogeneous measures, do not capture "individual differences in the direct encounter with media messages," but "assess exposure possibilities in the environment of individuals and try to show that such exposure opportunities are related to outcomes either across time or over geographic locations" (Liu & Hornik, p. 116). We distinguish three main types of measures. First, "media availability," measured by metrics such as circulation, ratings, or reach. Second, "campaign intensity," measured with metrics such as campaign expenditures, Gross Rating Points, or the number of passers-by of outdoor advertising. Third, media content data, which indicate an individual's media environment, useful for example in studies of "incidental exposure" (Hornik, 2016; Niederdeppe, 2014). Below we illustrate the application of these measures in media effects research.

Campaign tracking

An approach that uses ecological MCE measurement is campaign tracking which involves measuring media and campaign variables at an aggregate

Neijens, P., Araujo, T., Möller, J., & de Vreese, C. H. (2024). Ecological measures. In P. Neijens, T. Araujo, J. Möller, & C. H. de Vreese, *Measuring exposure and attention to media and communication. Solutions to wicked problems* (pp. 145-150). Amsterdam University Press. https://doi.org/10.5117/9789463723176_ch09

(site centric) level, including for instance, campaign expenditures, Gross Rating Points, site visits on the one hand, and campaign effect variables such as political, health and consumer cognitions, attitudes, and behavior on the other. Systematic tracking research measures these variables (usually) weekly, bi-weekly, or monthly with samples sizes of 100 or 200 respondents. Advertising tracking became popular in the late 1970s when Millward Brown started its tracking methodology and introduced the Advertising Awareness Index.

A challenge is the analysis of the "aggregated" data that requires modelling techniques as simplex models, autoregressive models, random effects models, multilevel models, latent growth curve models, or multivariate time series analysis (Berrington et al., 2006). Another issue is that the (weekly) effects of campaigns are generally speaking unspectacular and therefore difficult to detect. Some authors sigh that only big bursts of campaigning followed by a silence afterwards, can be noticed by the campaign evaluation studies. The — often — relatively small number of respondents in campaign evaluation surveys may also hinder the detection of campaign effects because of the large confidence intervals of the estimates. Furthermore, small sample sizes make it impossible to disaggregate impact scores to figures for different target groups, time periods, ad formats, and creative concepts, leaving the user with less informative averages (Merks-Benjaminsen, 2015).

Field experiments

A well-known example of the application of ecological measures are studies in which a "test market" (in which the campaign is running) is compared with a "control market" (Farquhar et al., 1990; Liu & Hornik, 2016; Luepker et al., 1994). These markets should be selected in such a way "that all the other variables that could affect the dependent variable remain approximately the same, or can be measured as covariates" (Tellis, 2004, p. 61).

The effects of exposure in field experiments include classic studies such as Berelson's analysis of what the strike of newspaper delivery meant for citizens suddenly not receiving their newspapers showing their news provision of public affairs changed as a result of the strike. A number of field experiments have been conducted in the realm of anti-smoking campaigns (Liu & Hornik, 2016) while another strand of research has been conducted on news and politics. Chaffee et al. (1990) showed how the availability of Korean television news in the Bay Area led to a higher uptake of this news by the community. Conroy-Krutz and Moehler (2015) randomly assigned passengers in mini vans in Uganda to exposure to radio talk shows and found that cross-cutting exposure to rival arguments was persuasive. Lelkes et al. (2020) in a nice design looking at differential access to internet broadband found that such access give rise to higher levels of partisan hostility.

In sum, these various field experiments, based on purposeful or natural variation, are interesting and convincing case studies of what changes in supply, or the opportunity structure of exposure to specific technologies or media content can do for different attitudinal and behavioral outcomes. However, while these studies have certain advantages, their quality also differs, as we can see in the next section.

9.2 Quality issues

Advantages of ecological measures are that memory problems that plague self-report are absent and that they mitigate concerns about reverse or reciprocal causality (Liu & Hornik, 2016; Slater, 2004). However, causality issues do exist: Comparing different time periods or different geographic locations may bring other (than campaign) variables into play. These factors include actions of competitors ("history"), seasonal and market trends ("maturation"), self-selection of respondents ("selection"), and methodological issues such as test-effects and statistical regression (Campbell & Stanley, 1963). For example, the choice of test markets may be influenced by the expectation that the specific market is more susceptible to change, or these markets may otherwise receive more attention from "campaign partners" such as healthcare providers or retailers, bringing the causality issue back on the table (Liu & Hornik, 2016).

The quality of the exogenous measures is not beyond doubt. Media expenditures are not necessarily a valid estimate of campaign intensity as market factors also determine media costs. Variables such as reach and circulation as indicators of media availability are self-reported and may suffer from their known problems (see Chapter 4).

Other problems are that ecological measures only measure the probability of exposure and not the actual exposure of an individual. Moreover, research designs with ecological measures do not allow to investigate subgroups and why a campaign did, or did not have the intended effect (Hornik, 2002a; Niederdeppe, 2014).

Liu and Hornik (2016) examined 80 studies that used ecological measures in tabaco research and found that only very few studies provided evidence on their validity. They suggested three general guidelines: first, match the research approach to the assumed paths of effects. "Aggregated measures may better capture supra-individual processes while they may be less sensitive to the discrete individual-level learning process" (p. 128). They also recommend to assess individual exposure differences, and to take into account the nature of the relevant content rather than just simple volume.

9.3 Concluding remarks

Ecological measures and field experiments are welcome approaches and designs in the pallet of the different ways to study exposure. They have certain properties that are favorable from a design perspective. But they are not without flaws, and feasibility is a major challenge in many cases. That said, they provide good real-life examples of the importance of exposure opportunities and differences in the supply of contents.

References

Berrington, A., Smith, P. W., & Sturgis, P. (2006). *An overview of methods for the analysis of panel data* (NCRM Methods Review Papers, NCRM/007). ESRC National Centre for Research Methods. https://eprints.ncrm.ac.uk/id/eprint/415/1/MethodsReviewPaperNCRM-007.pdf

Campbell, D. T., & Stanley, J. C. (1963). *Experimental and quasi-experimental designs for research*. Rand-McNally.

Chaffee, S. H., Nass, C. I., & Yang, S.-M. (1990). The bridging role of television in immigrant political socialization. *Human Communication Research*, *17*(2), 266–288. https://doi.org/10.1111/j.1468-2958.1990.tb00233.x

Conroy-Krutz, J., & Moehler, D. C. (2015). Moderation from bias: A field experiment on partisan media in a new democracy. *The Journal of Politics*, *77*(2), 575–587. https://doi.org/10.1086/680187

Farquhar, J. W., Fortmann, S. P., Flora, J. A., Taylor, C. B., Haskell, W. L., Williams, P. T., Maccoby, N., & Wood, P. D. (1990). Effects of community wide education on cardiovascular disease risk factors: The Stanford Five–City Project. *Journal of the American Medical Association*, *264*(3), 359–365. https://doi.org/10.1001/jama.1990.03450030083037

Hornik, R. C. (2002a). Public health communication: Making sense of contradictory evidence. In R. C. Hornik (Ed.), *Public health communication: Evidence for behavior change* (pp. 1–19). Erlbaum.

Hornik, R. (2016). Measuring campaign message exposure and public communication environment exposure: Some implications of the distinction in the context of social media. *Communication Methods and Measures*, *10*(2–3), 167–169. https://doi.org/10.1080/19312458.2016.1150976

Lelkes, Y. (2020). A bigger pie: The effects of high-speed internet on political behavior. *Journal of Computer-Mediated Communication*, *25*(3), 199–216. https://doi.org/10.1093/jcmc/zmaa002

Liu, J., & Hornik, R. (2016). Measuring exposure opportunities: Using exogenous measures in assessing effects of media exposure on smoking outcomes. *Communication Methods and Measures*, *10*(2–3), 115–134. https:/doi.org/10.1080/19312458.2016.1150442

Luepker, R. V., Murray, D. M., Jacobs Jr., D. R., Mittelmark, M. B., Bracht, N., Carlaw, R., Crow, R., Elmer, P., Finnegan, J., & Folsom, A. R. (1994). Community education for cardiovascular disease prevention: Risk factor changes in the Minnesota Heart Health Program. *American Journal of Public Health*, *84*(9), 1383–1393. https://doi.org/10.2105/AJPH.84.9.1383

Merks-Benjaminsen, J. (2015). *Online brand identity: The ultimate guide to designing your* (*digital*) *branding strategy*. Adfo Books.

Niederdeppe, J. (2014). Conceptual, empirical, and practical issues in developing valid measures of public communication campaign exposure. *Communication Methods and Measures*, *8*(2), 138–161. https://doi.org/10.1080/19312458.2014.903391

Slater, M. D. (2004). Operationalizing and analyzing exposure: The foundation of media effects research. *Journalism & Mass Communication Quarterly*, *81*(1), 168–183. https://doi.org/10.1177/107769900408100112

Tellis, G. J. (2004). *Effective advertising. Understanding when, how, and why advertising works*. Sage.

About the authors

Peter Neijens is an Honorary Fellow and Emeritus Professor at ASCoR. He held the Chair in Media and Persuasion at the University of Amsterdam until 2019. His research focuses on persuasive communication, in particular media use and effects in advertising and public opinion.

Theo Araujo is a Full Professor of Media, Organisations and Society, and Scientific Director of ASCoR. His research investigates the dynamic interplay between media and organizations, and what it means for society, with a special focus on trust and technology. His research has also a methodological focus on computational communication science and the implementation of large-scale data collection and analysis for communication research.

Judith Möller holds the chair of "Empirical Communication Research, especially Media Use and Social Media Effects" at the University of Hamburg and the Leibniz Institute for Media Research | Hans Bredow Institute (HBI). Before 2023 she was an Associate Professor of Political Communication at the Amsterdam School of Communication Research.

Claes H. de Vreese is a Distinguished University Professor of AI and Society with a special emphasis on media and democracy. He co-directs the Gravitation Program AlgoSoc and the AI, Media and Democracy Lab. He holds the chair in Political Communication at the Amsterdam School of Communication Research.

10. Recommendations

At the very core, the field of communication is concerned with detecting patterns in media and communication, understanding the processes that affect the production of, and the contingencies of the effects. Exposure therefore stands out as one of the crucial and unifying concepts: it is to some extent the "cradle of the field." On the one hand, if we fail to understand what it means to encounter media content, or are not able to measure it reliably and validly, the very foundations of the field are weak. On the other hand, thinking more deeply about the concept, offering reflections on developments, assessing weaknesses, and testing solutions, is also a genuine opportunity for the field. Either way, in our view, the exposure concept is a true core concept to keep paying attention to.

In this final chapter we offer recommendations for how to further improve the quality of media exposure measurement. Our recommendations concern theory, data collection, and future methodological research. But we start with a general, foundational recommendation.

1. Take the research question seriously

It is important to note that while we may sound prescriptive, a key message is that choices made by future scholars should very much depend on their lead question. The research question should guide considerations and choices, and from that perspective, an initial "ground zero" recommendation is to be explicit in the question formulation about the role that "exposure" has in the research endeavor. Is it conceptualized as a mediator, a dependent variable, or an independent variable? Is it conceptualized as part of research focusing on specific media, channels, platforms, or interactions? The more explicit these choices are, the better. Similarly, we also recommend being transparent about the considerations and choices, and — where appropriate and possible — to probe and test alternative conceptualizations, measures and/or analyses. It also seems both fair and necessary to distinguish between research that takes exposure as *the* central concept and research that relies

Neijens, P., Araujo, T., Möller, J., & de Vreese, C. H. (2024). Recommendations. In P. Neijens, T. Araujo, J. Möller, & C. H. de Vreese, *Measuring exposure and attention to media and communication. Solutions to wicked problems* (pp. 151-159). Amsterdam University Press. https://doi.org/10.5117/9789463723176_ch10

on exposure as *one* of several important concepts. The former may involve methodological studies or studies mapping (changes in) exposure patterns, whereas the latter may involve studies that try to assess, for example, the relative importance of exposure vis-a-vis other influences. In the latter case, there can be good and practical reasons for deploying shorter or more superficial exposure measures.

Recommendations for theory

2. Take media content and platform affordances into account

When developing and evaluating theories, concepts, and operationalizations related to media exposure, it is important to take into account the *content* of the medium and the *affordances* of the platform. In Chapter 2 we distinguished different levels and characteristics of the content (e.g., medium, program, title, genre, subject, frame, tone, valence), and the platform such as type of device (e.g., smartphone, tablet), application (e.g., social media, email), feature (e.g., reading, posting, chatting, gaming), interaction (active, passive), and message (e.g., mode, content). A way forward is to explicitly state the type of content, or affordance that is driving the causal mechanism being researched. This might be more or less independent of the specific media outlet individuals encounter it in. Thus, for example, rather than observing how much time a person spends on X or Facebook, the better question could be how much exposure this person had to a specific type of content, or type of affordance on those platforms. It will also allow researchers to focus their questions on the underlying principles rather than studying a specific platform (e.g., X).

De Vreese et al. (2017) offered guidelines for how to integrate content into survey questions, in an attempt to optimize linkage studies. In some cases, however, granular self-reports in combination with media content analyses might not be feasible, either because of researcher resources or because of the hybrid and complex exposures being tapped. On this topic, Guess et al. (2019) suggested that survey questions about social media use — which could be across platforms — should provide options for a wide range of activities, including those that occur less frequently (the "long tail"), and that general content probing questions, such as "being about politics," need specifics and anchors that define that content. In addition, new developments towards adaptive and quasi-automated content production (such as generative AI) will necessitate more information about the version of content that different individuals receive. For example, journalism might become more modular, tailored, and changing over time, making it less obvious to which version of content an individual was exposed.

3. Take context into account

Not only exposure, but also personal and social contexts that shape media exposure are relevant variables in theories on media uses and effects. Encounters that take place in more quiet surroundings (versus e.g., public, busy spaces), under the condition of single screens (versus multiple screens), as a primary activity (versus as one of a multi-tasking situation), as a single individual (versus with others, physically or online) are all potentially different encounters. This is especially relevant because technological developments point towards a further integration of the human body and technology. When that comes in the form of (portable) devices providing audio, wearable glasses, or virtual or augmented reality experiences, exposure may become more fluid. There is not a single answer as to how to incorporate these factors in the exposure concept. In some cases, the context in a study might be more or less constant or comparable for participants, in which case it could be less relevant for intra-participant observations, but in other cases it might vary in meaningful ways, and should be part of both the conceptualization and measurement. The challenge to capture context has become significantly larger through the unbundling of media content, for instance on recommender engines. This is true for the exposure to ads recommended by ad networks that display ads to targeted individuals, but also for news articles on news apps which are no longer ordered by the editorial board, or the interface of entertainment platforms which presents individual selections to each of their users. Since not only the selection of content, but also its potential effects are influenced by the context, we need to develop tools to capture this dynamic context. While this is certainly not easy, these tools will also provide us with the opportunity to test contextual effects in innovative and much more convincing ways than ever before.

When evaluating and applying theories and empirical findings, it is also important to take into account that the content, features, and use of media, platforms and applications may change over time. These changes can be technological (i.e., new features and affordances), substantive (e.g., new reporting style), or reflect changes in the media landscape (e.g., new outlets or formats). This is particularly pertinent if the leading research question is addressed in a *temporal* perspective, and for the assessment of research from different time periods. Formulating theories in terms of media affordances and media content (recommended above) will also allow for platforms to emerge, morph, or cease to exist in an over-time design.

4. Take physiological and mental processes during the exposure into account

Exposure is defined as "the extent to which audience members have encountered specific messages or classes of messages/media content" (Slater, 2004, p. 168). Although exposure as defined above is an extremely important variable, depending on the research question, physiological and mental responses during the encounter are relevant as well, as they may influence the effects of people's exposure to media. These responses include cognitive responses as attention, involvement, and engagement, affective responses as valence, arousal, and emotional responses, and experiences (see Chapters 2 and 4).

Recommendations for data collection

5. Take the measurement quality and representativeness of the data into account

It is absolutely necessary that studies examine and report the measurement quality of media exposure measures (such as validity and reliability), given a *precise* definition of what the measure is supposed to track — and under which conditions. Even in cases where we rely on established or previously tested measures, it is recommended to assess the quality of these in a new study.

As is pointed out, exposure measurement is context bound, and the precision, validity and reliability of a measure at one point in time is not a guarantee for its utility later. For example, measuring exposure to television news in the 1990s with much less supply, fewer channels, and fewer programs is different from measuring television news exposure in 2020 with a multitude of channels, web presences and news sharing via social media. Such developments beg re-assessing the quality of measures that might have performed well previously.

For the representativeness of the data, it is important to carefully consider the quality of respondent selection, (unit) non-response, and item non-response (compliance). Specifically for digital traces, representativeness can be threatened by platform coverage error, platform affordances error, and trace selection error (see Chapter 3).

6. Take all platforms into account

The trends in media usage are clear: less print, less traditional linear television, more digital, more streaming, more snippets of content across different platforms. This poses significant challenges to measuring exposure. Asking about watching a specific television channel or specific program might only capture a shrinking audience, and different versions of the same program might be available online. News organizations are reshaping their work flows to accommodate for the fragmented and diversified usage patterns: journalists are increasingly expected to deliver, for example, a two-minute item for a main television news show, a short piece for the website, a long read for behind the paywall, a photo for Insta, and a twenty-second video for TikTok, in addition to a mapping of how the stories were created and which choices were made to enhance transparency.

This has real implications for media exposure, which for many takes place on multiple platforms. This means that researchers need to take all of these platforms (and their characteristics) into account in order to fully understand an individual's media behavior. However, the degree to which how inclusive "all" is, again, truly depends on the leading question. If the question is about being exposed to "news about inflation" it would be important to consider the outlets, formats, and places in which such news might occur. However, a more narrow question about exposure to topic *x* on platform *y*, might lead to different choices and probing exposure on some platforms in more depth, while getting a more general and superficial estimate of exposure elsewhere.

7. Take the option to combine different measurement methods into account

Given the differences between, and the unique features of, different methods of measuring media exposure, combining methods in a single study is an attractive approach for a more complete understanding of an individual's media exposure. Again, this recommendation is qualified by the lead question. If this indeed evolves around getting a "complete" picture of an individual's exposure, an extensive battery of items and a combination of methods might be preferable. It is possible that the different exposure measures speak to different phases in the process of media effects. While eye tracking (Chapter 7), neurobiological measures (Chapter 8), or digital trace data (Chapter 5) can speak to the moment of exposure, self-reported measures are affected by what respondents can remember and have therefore processed. Differences between the various measures provide the opportunity to gain deeper insight into how and why information to which individuals are exposed is processed (see also suggestions in Chapters 4–8).

Recommendations for methodological research

8. Studies into validation and reliability needed

It is clear that all methods discussed in this book — retrospective and momentary self-reports, various methods to capture digital traces, data donations, observation methods, eye tracking, neurobiological and ecological measures — have their unique advantages. We cannot rule out one of them as inferior or suggest one of them as superior. There is only one thing to do: be conceptually clear, be explicit, be transparent, and continue improving measurement by further research, preferably embedded in systematic research programs.

It is impossible to present a comprehensive agenda for methodological research here, given the large number of (variants of) measurement methods, and the many issues involved. Methodological agendas are best drawn up for sub-areas. The chapters in this book highlight the unresolved issues that deserve a place on these agendas and provide inspiring examples of this type of research.

In research, attention to specific topics come and go. In communication science research, we would argue that the field cannot afford to lose attention for exposure. Constantly reflecting on the validity and reliability of measures in a changing technological landscape is a necessity.

9. Studies into non-response and non-compliance needed

All methods discussed in this book suffer from serious non-response and non-compliance issues. This is a problem with surveys, and even more so with intensive data collection methods such as logging of browsers and apps. Studies on the factors associated with these issues (see the corresponding chapters) should be continued, and approaches should be developed to improve representativity of media exposure studies.

10. Studies into merging datasets needed

Combining different measurement methods (see recommendation 7 above) poses conceptual and statistical issues that require more methodological research. This includes both "single source" data that captures all media use of an individual measured with different methods (e.g., digital traces data, self-reports), and data fusion, i.e., linking data sources by matching respondents on common variables, such as demographics, attitudes and behavior (see also Chapter 4).

Conclusions

These ten recommendations are meant to be specific and helpful in providing guidance about things to consider. They do not offer a quick decision tree on what to do, but rather a checklist of what to think about and make explicit. At some level, they might be considered frustrating due to the lack of clear-cut answers, but our ambition has been to recommend conscious decision making and explicit and transparent sharing of these choices.

In addition to the considerations around theorizing, data collection and (re-)boosting methodological research, a couple of broader industry and research community developments are worth mentioning. First, an enhanced dialogue between industrial research and academic research around exposure is welcomed. Today's institutionalized audience research, that necessarily makes use of different data sources and methods, has become extremely complex (see also Chapter 2). Given the importance of this data to industry, society, and policy makers, independent assessments of their quality are necessary and must be made possible.

Second, research on media exposure also has work to do in terms of ethical guidelines and privacy: lengthy survey questionnaires, laboratory studies, and data donation studies, all in different ways, are rather taxing for participants. Establishing good practices in terms of minimizing participant burden is a point of attention, in addition to more recent requirements about enhanced data management and privacy aspects forcing researchers and institutions to improve their infrastructure and processes for data collection, management, analysis, and archiving.

Third, new European Union regulations such as the Digital Services Act open new opportunities for media exposure research: Very large online platform (VLOPs, defined as having >45 million users in Europe) are compelled to share certain data with the research community. This will open new avenues for media exposure research and potentially offer more insight into thus far closed platforms.

Fourth, the ongoing movement towards enhanced transparency and open scholarship practices, should also lead researchers engaging with the media exposure concept to become more open and sharing in terms of measurements and analyses (see also Chapter 8). This should further improve the development of and reflection on exposure measures.

Long live exposure

The exposure concept comes close to a raison d'être for the field of communication. It forms the basis for many core questions. At the same time, it remains a constant challenge. Self-reported exposure measures, even decades ago with a media landscape consisting of just one or a few television channels, a handful national radio stations, and a similar number of leading newspapers led to questionable estimates. Since then, the media landscape has undergone tremendous changes in multiple ways. Each of these have arguably both worsened the case for exposure measures but also offered new data sources and ways of conceptualizing media and communication exposure.

The concept is criticized and discussed in oftentimes nihilistic terms. However, we remain "glass half full" researchers: our toolkit has expanded enormously in the past, say, two decades and new opportunities have arisen for conceptualizing, measuring, and modelling exposure. We hope our overview is useful in providing a range of things to consider. There is no quick fix, magic bullet or easy solution. But there is an invitation to think twice about the concept before deploying it, to be explicit about assumptions, choices and alternatives, to be transparent about the implications, and to engage in (programmatic) research on the exposure concept itself.

References

de Vreese, C. H., Boukes, M., Schuck, A., Vliegenthart, R., Bos, L., & Lelkes, Y. (2017). Linking survey and media content data: Opportunities, considerations, and pitfalls. *Communication Methods and Measures*, *11*(4), 221–244. https://doi.org/10.1080/19312458.2017.1380175

Guess, A., Munger, K., Nagler, J., & Tucker, J. (2019). How accurate are survey responses on social media and politics?. *Political Communication*, *36*(2), 241–258. https//doi.org/10.1080/10584609.2018.1504840

Slater, M. D. (2004). Operationalizing and analyzing exposure: The foundation of media effects research. *Journalism & Mass Communication Quarterly, 81*(1), 168–183. https://doi.org/10.1177/107769900408100112

About the authors

Peter Neijens is an Honorary Fellow and Emeritus Professor at ASCoR. He held the Chair in Media and Persuasion at the University of Amsterdam until 2019. His research focuses on persuasive communication, in particular media use and effects in advertising and public opinion.

Theo Araujo is a Full Professor of Media, Organisations and Society, and Scientific Director of ASCoR. His research investigates the dynamic interplay between media and organizations, and what it means for society, with a special focus on trust and technology. His research has also a methodological focus on computational communication science and the implementation of large-scale data collection and analysis for communication research.

Judith Möller holds the chair of "Empirical Communication Research, especially Media Use and Social Media Effects" at the University of Hamburg and the Leibniz Institute for Media Research | Hans Bredow Institute (HBI). Before 2023 she was an Associate Professor of Political Communication at the Amsterdam School of Communication Research.

Claes H. de Vreese is a Distinguished University Professor of AI and Society with a special emphasis on media and democracy. He co-directs the Gravitation Program AlgoSoc and the AI, Media and Democracy Lab. He holds the chair in Political Communication at the Amsterdam School of Communication Research.

Index

Accuracy 34, 46, 57, 59, 64-66, 122
Affective
aspects 68
behavior 90
dimensions 131
experiences 134
impressions 57
processes/processing 18, 116
responses 15, 18, 67, 130, 132, 136, 154
Affordances 33, 37, 152-155
AIR, *see also* Average Issue Readership 24
Algorithmic errors 96
Anchoring 52
Anchors 50-52, 152
Answer scales, *see also* response scales 49, 59
AOIs, *see also* Areas of Interest 113-115, 117, 119-122
AOI time 113-115, 117, 122
API, *see also* Application Programming Interface 88, 89, 94
App activity 90
Application Programming Interface, *see also* API 88
Areas of Interest, *see also* AOIs 113, 115
Arousal 15, 18, 24, 67, 69, 130-131, 133, 136, 137, 154
Associated score 24
Audience
composition 20, 23
duplication 23, 24
size 23
Audiometer 21
Augmented reality
ethnography 108
experiences 153
Average (contact) frequency 23-25
Average Issue Readership, *see also* AIR 24

Beacons, *see also* grabbers 88, 89
Between-person 35, 62, 66
Bots, *see also* web crawlers 88, 89, 94
Bottom-up processing 16, 118

Campaign tracking 145
Census data, *see also* platform centric, *see also* server centric, *see also* site centric 22, 23, 87, 89
Channel levels 19
Circulation 24, 26, 146, 147
Click-through 23, 24, 120
Cognitive responses 67, 133, 136, 154,
Communication power 21
Compliance 37, 41, 54, 63-65, 90, 97, 98, 154, 156
Contact frequency 23
Contact quality 20
Content analysis 19, 36, 42, 43, 46, 49, 53, 123, 138, 152
Context 15, 41, 49, 66, 70, 105, 116, 122, 153, 154
Coverage errors 37, 155
Cross media measurement 21, 22
Cumulative measures 23
Currency 12, 20

Data donation 64, 87-89, 91, 92, 94, 95-99, 157
Data Download Package, *see also* DDP 91
Data linking, *see* linkage
Day Construction Method 51
DDP, *see also* Data Download Package 91, 92, 97, 98
Diary 21, 22, 35, 41, 51, 54, 65, 109
Digital ethnography 105, 108, 109
Digital Services Act 157
DSA, *see* Digital Services Act
Duplication, *see* audience duplication
Duration 41, 43, 44, 58, 59, 61-64, 66, 95, 96, 99, 108, 114, 115, 116, 122
Dwell time 115-117

EARS, *see also* Effortless Assessment of Risk States 90
Ecological Momentary Assessment, *see also* Experience Sampling, *see also* In situ Measurement, *see also* Real-Time Response Measurement 54
EEG 129, 130, 135
Effective reach 25
Effortless Assessment of Risk States, *see also* EARS 90
Elaboration 16, 56
EMA, *see also* Ecological Momentary Assessment 54, 63-65
EMM, *see also* Eurisko Media Monitor 90
Emotional responses 15, 18, 69, 154
Engagement 15, 17, 18, 21, 24, 25, 67-70, 87, 88, 94, 134, 154
ESF, *see also* experience sampling form 54
ESM, *see also* experience sampling method, *see also* ecological momentary assessment, *see also* experience sampling, *see also* real-time response measurement 54, 63-66
Ethical concerns/ issues 53, 93, 99, 109, 157
Eurisko Media Monitor, *see also* EMM 90
Exogenous measures, *see also* ecological measures
Experience sampling, *see also* Ecological Momentary Assessment, *see also* In situ Measurement, *see also* Real-Time Response Measurement 54, 66
Experience sampling form, *see also* ESF 54

Exposure states 17, 18
Extraction errors 96
Eye trackers 116-118, 123

Facial Electromyography, *see also* fEMG 131
Fatigue effect 66, 90
fEMG, *see also* facial Electromyography 129, 131, 132
Fixation
  count 115
  duration 115
  frequency 115
Flow 17, 134
fMRI, *see also* functional Magnetic Resonance Imaging 12, 129, 130, 133-135, 137
fNIRS, *see also* functional Near Infrared Spectroscopy 135
Focal attention 16
Foils, *see also* ringers 44
Frame error 36
Fraud 94
Frequency list technique 48
Functional magnetic resonance imaging, *see also* fMRI 133
Functional near infrared spectroscopy, *see also* fNIRS 135
Fusion, *see also* merging 22, 157

Gaze 67, 108, 115, 118, 132
GDPR, *see also* General Data Protection Regulation 91
General Data Protection Regulation, *see also* GDPR 91
Google Analytics 88
Grabbers, *see also* beacons 88
Gross measures 23
Gross Rating Points, *see also* GRPs 23, 25, 26, 145, 146
Gross reach 23, 25
GRPs, *see also* Gross Rating Points 23, 25

Heart rate 129, 130, 133
Help, *see also* instruction 42, 43, 50, 70
Hertz, *see also* sampling interval, *see also* sampling frequency 115, 121
Hooperratings 21
Household meter 21, 89, 90

Impressions 23, 25
In situ Measurement, *see also* Ecological Momentary Assessment, *see also* Experience sampling, *see also* Real-Time Response Measurement 41, 43-45, 53, 65
Incentive 95, 96
Information processing 15, 56, 129, 130, 133, 137
Information retrieval measures 36
Institutionalized media audiences 15, 20, 21, 157
Internal consistency 34, 106
Involvement 15-18, 21, 24, 45, 57, 66, 67, 68, 154

JIC, *see* Joint Industry Committee
Joint Industry Committee 20

Liking 68, 93
Linkage 19, 42, 43, 52, 53, 70, 152
List measures 43, 47, 48
Log-based measures 63, 64

Market share 23
MCE (Medium and Communication Exposure) 11, 13, 33, 37, 145
Metrics 15, 20-24, 87, 88, 90, 91, 113, 115, 116, 121, 122, 145
Measurement errors 33, 37, 44, 53, 93, 96, 121
Measurement unit 43, 44, 55
Media
  effects 19
  experiences 18, 70, 123
  reception 18
  use 18
Medium type 19, 42
Mental processes 15, 154
Mental responses 24, 41, 154
Merging, *see also* fusion 157
Message processing 16
Momentary measurement, *see also* Ecological Momentary Assessment, *see also* Experience sampling, *see also* In situ measurement, *see also* Real-Time Response Measurement 54
Motivational state 18
Multitasking 12, 120, 123
MVPA (Multivariate Pattern Analysis) 134

N=1, *see also* within-person 36, 62
Noted scores 25
NPP (Neurophysiological perspective) 130, 133

Obtrusive, obtrusiveness 55, 109
Operationalizations 9, 12, 16, 23, 34, 36, 152
OTS (Opportunity To See) 16, 25
Overestimation, *see also* overreport 35, 52, 57, 59, 61, 65
Overlap 22, 25
Overreport, *see also* overestimation 49, 61, 62

Page views 25
Pay per click 94
PANAS 69
People meter 22, 61, 89, 90
Personal involvement inventory 68
Physiological 15, 130-132, 135-137, 154
Platform affordances error 33, 37, 152, 155
Platform centric, *see also* site centric, *see also* server centric, *see also* census data 23, 87, 88
Plugins 91, 97
Portable devices 153
Portable people meters 89, 90

Positive and Negative Affect effect Schedule, *see* PANAS
PPM 90
Preattention 16
Precision 36, 92, 118, 121, 122, 154
Program list technique 47, 48
Prompt 64

Questionable Research Practice (QRP) 135

Ratings 24, 25
Reach 23-25
Read most 25
Readership 23, 24
Reaction times 67
Real-time Response Measurement (RTR), *see also* Ecological Momentary Assessment, *see also* Experience Sampling, *see also* In situ Measurement 54
Recall 42-70
Recency 44, 45, 59
Recent reading 23, 25
Recognition 24, 25, 42, 43
Reference period 43-46, 51
Regions Of Interest (ROI) 134
Reliability 33-36
Representation error 37, 97
Response scales, *see also* answer scales 49, 50
RFID 88, 89
Ringers, *see also* foils 44
Router meter 22, 89, 90

Saccade 115
Sampling errors 33
Sampling frequency, *see also* sampling interval, *see also* Hertz 115, 121
Sampling interval, *see also* sampling frequency, *see also* Hertz 115, 121
Scanpaths 115
Screen time 20, 25, 91, 92, 96, 98, 106
Screen use 25
Selectivity 16, 18, 25
Self-reflexive state 17
Sensitivity 34, 95, 98
Server centric, *see also* census data, *see also* site centric, *see also* platform centric 23, 87
Server log data 62
Set-top box 89
Share of market, *see also* share of voice 23
Share of voice, *see also* share of market 23
Sharing 68, 70
Silo approach 22
Single source 22, 157
Site centric, *see also* census data, *see also* server centric, *see also* platform centric 23, 87-89, 146
Skin conductance 12, 129-131, 136
Social desirability 44, 47, 56, 60, 61, 66, 118, 119, 132
Specific Issue Readership 23
Stimulus time 114, 115
Subscriptions 26

Target Rating Points 26
TED-On 93
Telephone coincidentals 53
Think aloud, *see also* thought listing 12, 55, 56, 66, 67
Thought listing, *see also* think aloud 17, 66
Through-The-Book method 23, 42
Time spent 25, 44, 62, 65, 66, 68
Time To First Fixation 115, 121
Top-down processing 16
Total survey error 33
Total survey quality 33
Trace selection error 37, 155
Transported state 17
Triangulation 96, 113, 116, 122
TRPs, see Target Rating Points
TSE, see Total Survey Error
TSM, *see also* time spent 62
TTB, *see* Through-The-Book method
TTFF, *see* Time To First Fixation

Unique visitors 23, 26
Unit of measurement 43, 44, 55
Unpacking effect 46
User-centric 23, 87-89
User selection error 37
User tracking 87, 89-91, 95

Valence 15, 18, 20, 24, 67, 69, 131, 152, 154
Validity 34, 35
Vehicles 19, 21, 24
Views 25, 88, 94, 132
Virtual reality ethnography 108
VLOPs, Very Large Online Platforms 157
Volume 43-45, 64, 148

Waste 26
Wearables 89, 90, 137, 153
Web crawlers, *see also* bots 88, 89
Web harvesting 88
Web Historian 92
Web scraping 88, 89
Within-person, *see also* N=1 35, 62, 66

Valid and reliable measurement of media and communication exposure is crucial for communication science, psychology, political science, sociology, pedagogy, economics, and law, and the practitioners in media, communication, and information. At the same time, this is a wicked problem for which there are no simple solutions. That was never the case, but in today's digital and abundant media landscape it is even more difficult. This book discusses the ways in which media and communication exposure can be conceptualized, operationalized, and measured. Methods examined include self-reports, recall, recognition, ecological momentary assessment, think-aloud, digital traces, data donation, human observation, eye-tracking, EEG, fMRI, heart rate, and skin conductance, their pros and cons, complexities, and performance. The book concludes with recommendations for the application and further development of these methods. It also includes an extensive bibliography — with references to in-depth insights into specific aspects of media exposure measurement.

"This book will be an invaluable resource for training graduate students and for exploring research design alternatives by faculty and industry researchers in communication, media psychology, and allied fields."
Michael Slater, Social and Behavioral Sciences Distinguished Professor School of Communication, The Ohio State University, USA

"I am confident that this book will provide an excellent resource for many, from students to experienced experts in the field, and that it will instill much-needed future discussions and research on the conceptualization and measurement of this core construct."
Veronika Karnowski, Chair of Media Communication, Chemnitz University of Technology, Germany

ISBN: 978-94-6372-317-6

9 789463 723176

AUP.nl