METHODEN UND FORSCHUNGSLOGIK
DER KOMMUNIKATIONSWISSENSCHAFT

Christina Peter / Teresa K. Naab / Rinaldo Kühne (Eds.)

Measuring Media Use and Exposure

Recent Developments and Challenges

HERBERT VON HALEM VERLAG

VERONIKA KARNOWSKI / TERESA K. NAAB /
DANIELA SCHLÜTZ

On the Challenges of Measuring Mobile Social Media Use: Explaining Differences Between Data from Surveys and Mobile Experience Sampling

Quantifying media use is a constant challenge in communication research. With media exposure being a central construct in many areas of research, valid and reliable measurement is particularly crucial (DE VREESE/NEIJENS 2016; FISHBEIN/HORNIK 2008; SLATER 2004). Standardized retrospective surveys are most common to capture media use (HA et al. 2015). However, they demand a great deal of respondents, especially regarding their ability to remember and generalize media behavior, because self-reports mostly focus on overall reports instead of specific answers on recent behavior.

A possibility to address this shortcoming of retrospective surveys (besides applying observational methods, DE VREESE/NEIJENS 2016) is to collect data by means of the Experience Sampling Method (ESM): ESM repeatedly collects data from participants at multiple points in time. Participants are asked to respond with as little time lag between exposure and self-report as possible and thus produce in-situ data with little need to remember and to generalize.

In this chapter, we first outline causes of measurement errors in retrospective self-reports. Thereby, we specifically regard the measurement of mobile social media behaviors. We then introduce the Mobile Experience Sampling Method (MESM) in general, its special benefits to measuring mobile social media use, and its potential to overcome the measurement challenge presented by retrospective self-reports. Based on this, we deduce potential factors explaining differences between data obtained by retrospective surveys and MESM, respectively. These factors are then brought to a test in an empirical study. We compare both methods for measuring the duration of YouTube, Facebook, and WhatsApp usage episodes and analyze determinants of the difference between the duration values obtained by these methods.

1. Measurement errors in self-reports

The survey method is valid for measuring subjective conditions (like attitudes, emotions, cognitions) provided these are conscious and reproducible. It is less valid, however, for measuring behavior (SCHERPENZEEL/SARIS 1997), as retrospective interviews demand a great deal of respondents regarding their ability to remember and to generalize their behavior (BRADBURN/RIPS/SHEVELL 1987; NISBETT/WILSON 1977). Empirical research has repeatedly documented measurement errors when survey answers are compared to behavioral data like log or provider data (BOASE/LING 2013; JUNCO 2013; KOBAYASHI/BOASE 2012; PRIOR 2009b; SCHARKOW 2016; VANDEN ABEELE/BEULLENS/ROE 2013).

Prior (2009a) assumes that reporting errors result from unrealistic demands on respondents' cognitive abilities. Research on cognitive aspects of survey methodology suggests that questions activate a multi-level cognitive process consisting of (at least) five steps: The interviewee has to understand the question (step 1: comprehension), recall the relevant behavior and/or cognitions (step 2: information retrieval), make judgments (like inferences and estimations) concerning these behaviors and cognitions (step 3: judgment), adapt her or his answer to fit the response format (step 4: reporting), and edit the answer for reasons of social desirability or self-presentation (step 5: adjustment; SCHWARZ/OYSERMAN 2001; SCHWARZ/OYSERMAN/PEYTCHEVA 2010; TOURANGEAU/RIPS/RASINSKI 2000). This process implies that cognitions are always accessible, that people know what they do, and that they can report on their behavior accurately, even if they may not be willing to do so (SCHWARZ/OYSERMAN 2001). If respondents undergo this process (at least steps 1 to 4) diligently, their estimates are likely to be an accurate representation of their behavior. This may not always be the case, though. Rather, in many instances, respondents may give the most convenient answer instead of the most accurate one. This answering heuristic has been termed ›satisficing‹ (KROSNICK 1991). Respondents may shortcut the cognitive process by running through all steps without any specific effort (weak satisficing) or by omitting some steps altogether (strong satisficing). A person's propensity to satisfice varies with the difficulty of the question, his or her cognitive ability, motivation to perform the task, and willingness to state the correct answer (SCHWARZ/OYSERMAN 2001). If, on the other hand, respondents answer thoughtfully and accurately, they are said to ›optimize‹.

2. Specific challenges in the measurement of mobile social media use

The methodological challenges of measuring media exposure intensify when it comes to mobile social media. Their use can vary from bidirectional to rather unidirectional use, from synchronous to asynchronous, and from text over voice to video. It can spread over a variety of platforms of all kinds of professional and user-generated content. Hence, it comprises an unprecedented breadth of patterns of use. Thus, retrospective self-reports are particularly demanding (NIEDERDEPPE 2016; DE VREESE/NEIJENS 2016).

Moreover, the usage of mobile social media is both ubiquitous and highly volatile. For decades, media use was tied to stable situational contexts. Most media (except for very few devices like newspapers or transistor radios) were bound to specific locational settings (QUANDT/VON PAPE 2010). These stable spatial surroundings went along with very little variation in social contexts. In the era of desktop computers, variation in online usage situations was little, with usage normally being restricted to either at home or at the work place (HARGITTAI/HINNANT 2008). The introduction of wireless networks and portable computers extended the range of online media use (FELDMANN 2005; HAMPTON/LIVIO/SESSIONS GOULET 2010). The spread of 3G networks and smartphones finally led to truly ubiquitous and mobile online media use (HESS 2007; WESTLUND 2008). Henceforth, online media use, including social media use, started to penetrate even the smallest niches in our everyday lives, thus not only introducing social media use to a theoretically unlimited array of situational contexts but also bringing along situational characteristics as a new set of factors influencing media use (KARNOWSKI/JANDURA 2014; STRUCKMANN/KARNOWSKI 2016). From a methodological point of view this dramatically increased breadth of situational contexts is a problem: If media use co-varies with situational contexts and these contexts become more versatile and multifaceted, it gets more and more difficult to estimate average media use or to recall a ›typical‹ usage situation correctly in retrospect.

The diffusion of mobile online media use into the niches of our everyday lives not only came along with a broad array of situational contexts but also with a dramatically increased frequency of sometimes very short usage episodes, a phenomenon also termed as being permanently online and permanently connected (VORDERER/KRÖMER/SCHNEIDER 2016). This acceleration of media use is part of the broader phenomenon of social acceleration (ROSA 2015). Ubiquitous media access intensified this already existing trend, however: first manifesting itself in the taken for grantedness of mobile communication services throughout our everyday lives (LING 2012) and nowadays cumulating in the emergence of mobile social media apps like Snapchat, with the ephemerality of the very moment being a constituent characteristic of the service (BAYER et al. 2015). Similar to the increase in situational contexts, we assume this volatility or even ephemerality of mobile social media use to impede accurate recall in retrospective surveys. Concentrating on the specific case of mobile phone calling and texting behaviors, Boase and Ling (2013) already found such ubiquitous high frequency and low duration behaviors to be prone to reporting bias. The authors called for alternative methodological approaches to measure mobile media use.

3. (Mobile) Experience Sampling Method

The Experience Sampling Method (ESM) is a method of data collection in which respondents repeatedly report on behavior, cognitions, emotions, and situational aspects of their surroundings. Data are collected in natural settings and across situations over a certain time span (LARSON/CSIKSZENTMIHALYI 1983). At multiple times, respondents receive a signal and are asked to answer a short written questionnaire – called Experience Sampling Form (ESF) – on their current situation with as little delay as possible (CONNER/BLISS-MOREAU 2006). Respondents are either notified to answer an ESF at random points in time or in predetermined regular time intervals. Alternatively, in event-contingent sampling, participants report during or immediately after predetermined events (SCOLLON/KIM-PRIETO 2003), for example after having made a mobile phone call (COHEN/LEMISH 2003).

In communication research, ESM has only been scarcely employed (but see SCHLÜTZ 2002), probably mostly because of its drawbacks. In its beginning, the method was rather cumbersome for participants, as they had to carry along a pager to receive the signal as well as paper ESFs to respond in time. The traditional ESM did not provide the researcher with any kind of control on whether the ESF had been completed in-situ or delayed and thus with no advantage over a retrospective survey. This is no longer true for ESM's up-to-date mobile version called Mobile Experience Sampling (MESM; BOLGER/DAVIS/RAFAELI 2003): Participants are notified and interviewed via smartphone, thus no additional signaling device is necessary. The ESF is directly administered as a reply text message (e.g., COHEN/BOWMAN/LANCASTER 2015), by a specific app (e.g., BOASE/KOBAYASHI 2012), or as a web-based, mobile-optimized online questionnaire (e.g., KARNOWSKI et al. 2017). Both app and mobile-optimized questionnaire offer the technical benefits of online surveys like filters, multi-media components, additional data types (like photographs, geodata, or other forms of data traces), and time stamps (e.g., BRANDT/WEISS/KLEMMER 2007; PALEN/SALZMAN 2002). Time stamps allow to control for (or even impede) time lacks between prompting and answering.

Despite the assumed benefits to validity and reliability, MESM studies also face challenges known to most panel research including biased samples because of panel mortality, panel bias, decreasing compliance, and reactivity of the method because of heightened sensitivity (CSIKSZENTMIHALYI/LARSON 1987; see also SCHERER/NAAB 2013). However, research into the method indicates that increased self-awareness during the period of study should not be cause for general concern: Split-half experiments show that although answering delay increases slightly, the use of the scale does not vary significantly between the first and the second week of the study (SCHLÜTZ 2002; SCHLÜTZ/SCHERER 2001). Another methodological problem concerns communication research in particular: When investigating the use of media devices and content which the respondents use with low frequency or short duration, a major proportion of the situational measurements might not capture situations with media use but other pastimes. This is an interesting result, but it may increase costs and time frame of the study. This problem can be circumvented by using event-contingent instead of random sampling. Additionally, there might be expenses for the participants for receiving text messages and accessing the Internet. Thus, incentives to ensure compliance are necessary. Finally, MESM-studies pose challenges with regard to data security and privacy protection that have to be considered (BOUWMAN et al. 2013). A more practical disadvantage of MESM studies is the mandatory availability of a smartphone with permanent Internet access for each subject. For certain research questions, this imposes limitations for sampling and, consequently, representativeness and generalizability of the results.

4. Explaining differences between measures obtained using retrospective and in-situ self-reports

There are two different strategies to elicit self-reports of behaviors (step 2 of the above-mentioned answering process): Either participants are asked to recall ›typical/usual‹ media behavior or media behavior in a ›recent/past‹ specific time span. These approaches are called frequency and recency method, respectively (CHANG/KROSNICK 2003; PRIOR 2009a). Frequency and recency questions are likely to yield different results: After a certain time, episodic memories become too few to aid reports, and respondents rely on semantic memory that consists of generalizations instead of knowledge on particular experiences. In contrast, self-reports on specific, more recent, and shorter episodes are assumed to be comparably less biased than ›typical‹ ones (KAHLOR et al. 2003; LEE/HORNIK/HENNESSY 2008). This might result in differences between retrospective data, which asks for typical behavior, and in-situ measurement of media use, which asks for recent, specific behavior (ROBINSON/CLORE 2002).

In MESM studies, participants are asked to assess their momentary behavior or their behavior during a very recent time span like 30 minutes or one hour. Respondents are instructed to respond with as little delay as possible. Thus, MESM applies repeated recency measures. Compared to retrospective surveys applying the frequency method, MESM offers advantages for both external and internal validity because data are gathered in-situ and influences of the method are rare (see above and CSIKSZENTMIHALYI/LARSON 1987; KUBEY/CSIKSZENTMIHALY 1990; SCHLÜTZ 2002; SCHLÜTZ/SCHERER 2001).

In addition to lack of memory, generalizations made over a longer time span across situations may be invalid. In retrospective surveys with frequency questions respondents need to recall several distant events in time and to aggregate these events over a longer interval in order to translate their behavior into an answer (step 3 of the answering process). »In most cases they rely on extensive inferences and estimation strategies to arrive at an answer« (SCHWARZ 2007: 13). Such aggregation is prone to measurement errors, because past behavior is more difficult to account for than more recent behavior (KROSNICK 1991; PRIOR 2009a; see also BERNARD et al. 1984). This is especially true when the behavior in question lacks inter-situational stability (LEE/HORNIK/HENNESSY 2008; see below). In order to avoid miscalculations, the recency method may be employed, because self-reports on more recent, specific episodes are assumed to be comparably less biased (KAHLOR et al. 2003; LEE/HORNIK/HENNESSY 2008). Here, respondents are asked to recall a specific and more confined period, for instance the last day, and to report on their media use accordingly (e.g., BOASE/LING 2013). This does not demand averaging by the respondents. Unfortunately, this method is prone to errors as well as particular past behavior media use might be atypical compared to average use (CHANG/KROSNICK 2003). This is especially pressing when measuring constructs that assumedly vary across situations (SCHNAUBER 2017).

MESM provides an advantage in this regard, too. A single report in one ESF may paint an atypical picture when the respondent's media use is highly irregular and related to a very specific situation. However, MESM is a longitudinal approach, and participants repeatedly state their behavior across a multitude of randomly chosen situations. Every time, they answer on the specific, recent situation with no need to recall a long time span or to aggregate several different use episodes. Thus, the approach employs the advantage of the recency method in that respondents only have to report on their behavior in a specific situation. This increases ecological validity and generalizability of results. Additionally, information about average usage behavior can be computed by aggregating each participant's situational information to trans-situational information. Thus, information usually pertained by the frequency method can be computed.

Given these differences in methods, it is highly likely that data on media use gathered by a retrospective survey and MESM, respectively, differ. This has been shown for mobile social media use (NAAB/KARNOWSKI/SCHLÜTZ 2018): The authors present data that several characteristics (i.e., duration and frequency of a usage episode, habit, elaboration, and gratifications) of the use of the mobile social media platforms Facebook, WhatsApp, and YouTube vary between retrospective survey and mobile experience sampling measurement. This paper sets out to explore factors influencing these differences. In doing so we will focus on one specific measure of media use, duration of a single usage episode. This is a quite precise measure of exposure (as compared to frequency of use) used in many communication studies. The extent of media exposure is a prerequisite for subsequent attention, comprehension, and retention (SLATER 2004: 168).

The different approaches described above are assumed to differ in outcome as a consequence of greater demands to retrieve and to generalize past behavior in retrospective surveys. Following extant empirical evidence, these differences should vary as a function of specific characteristics of the examined media use, namely behavioral frequency, habit strength and context stability, involvement, and social desirability.

4.1 *Influence of behavioral frequency*

The accessibility of past behavior is especially questionable when it comes to high-frequency behavior. People are unlikely to have detailed representations of frequent and closely related behaviors. Instead, the numerous instances »blend into one global knowledge-like representation that lacks specific time or location markers« (SCHWARZ/OYSERMAN 2001: 136-137). Consequently, respondents have been found unable to distinguish and retrieve individual episodes (ROBINSON/CLORE 2002) in order to give valid answers representing their ›true‹ media behavior. Especially reports on mundane behaviors are prone to this error, while »more distinct events, in terms of intensity, emotionality, unusualness, or personal significance« (REIS/GABLE 2000: 196) tend to be recalled better (e.g., BOASE/LING 2013 on voice calls and text messages with mobile phones). As has been outlined above, mobile media use comes with an unprecedented high frequency compared to traditional media use. Additionally, single usage episodes have a low duration (VORDERER 2015). Generally, people tend to underestimate the occurrence of high-frequency behaviors and to overestimate the occurrence of low-frequency behaviors (SCHWARZ 2007). This should result in different (averaged) estimations in retrospective self-reports compared to more recent reports. Thus, we assume:

H1: The difference between retrospective and in-situ measurement of media use duration varies as a function of the frequency of the media use in question.

4.2 *Influence of habit strength and context stability*

The more varied and inconsistent over time a media behavior is, the more demanding it will be for respondents to recall instances of this behavior and translate these into an answer to a frequency question regarding typical behavior. Instead of computing a mean value, respondents are more likely to rely on extensive inference strategies (SCHWARZ 2007). Regular behavior might be estimated correctly with these strategies, but irregular, less consistent media use is prone to greater measurement error. Accordingly, it has been shown that reliability of self-reports on past behavior increases with the stability of the behavior (ROSS 1989; SCHWARZ/OYSERMAN 2001; also LEE/HU/TOH 2000).

Behavioral stability can be understood with regard to the construct of habit strength. When faced with repetitive situations, people do not need to deliberate on potential behavior but activate a mental representation of the recurring situation and perform the stored behavior. This behavior is called a habit (LAROSE 2010; NAAB/SCHNAUBER 2016). The habit is triggered by repetitive situational cues like place, social circumstances, or internal conditions (DANNER/AARTS/VRIES 2008). As habitual behavior occurs with little variation, habitualization of media use should facilitate inferences and thus limit the discrepancies between retrospective and in-situ measurement. Hence, we deduce:

H2: The difference between retrospective and in-situ measurement of media use duration varies as a function of habit strength of the media use in question.

Furthermore, the stability of the context of media use may influence the accessibility of answers and the need for heuristic self-reports. This is of specific significance for the measurement of mobile social media use. Since – as outlined above – the situational context is widely inconsistent across usage situations, respondents cannot rely on deducing estimates of media behavior from estimates of context cues. For example, an answer to the question ›How often did you watch the news last week?‹ could be derived from the estimation how often one was at home during the evening last week. However, such a heuristic is less helpful when one not only watches the news at home on television but also on the smartphone not bound to a particular location and a certain time of linear broadcasting. Based on this, we assume:

H3: The difference between retrospective and in-situ measurement of media use duration varies as a function of context stability of the media use in question.

4.3 *Influences of involvement*

Valkenburg and Peter (2013: 200) argue that motivational factors like involvement with the topic that predict attention to a particular message may be confounded with awareness and recall of the message (SCHWARZ/OYSERMAN 2001). Involvement describes that people cognitively, affectively, conatively, and motivationally engage with an issue (WIRTH 2006), that is involvement is the state of felt importance of a specific media content. Involvement, in turn, may bias self-reports of exposure: The more people are involved in media use, the easier it should be for them to remember the episode and the more precise should their estimate be. Hence, we postulate:

H4: The difference between retrospective and in-situ measurement of media use duration varies as a function of the involvement in the media use in question.

4.4 *Influence of social desirability*

Biases can be caused by questions and item wordings that direct to issues on which respondents hesitate to give accurate answers because they perceive their answer as socially inappropriate and fear disapprovement by others (social desirability bias, PAULHUS et al. 2003). A meta-analysis by Scherpenzeel and Saris (1997) shows lower validity and reliability of questions that are sensitive to allegedly social desirable answers. Apart from a characteristic of an issue or a question, social desirability is also considered a trait of participants understood as »a tendency on the part of respondents to give favorable impressions of themselves« (DEMAIO 1984: 276). Some people are more likely to adapt their answers with regard to a presumably socially tolerable norm (CROWNE/MARLOWE 1960). However, there is evidence that frequent self-reports in short time intervals, as given in diary or ESM studies, limit the effect of social desirability, while retrospective evaluations are more prone to these effects (CARP/CARP 2007). However, this should only be the case for behavior that is perceived as less social desirable by the respondents. Hence, we postulate:

H5: The difference between retrospective and in-situ measurement of media use duration varies as a function of a tendency to give socially desirable answers with regard to less socially desirable media behavior.

5. Method

5.1 *Overview*

The study consists of two steps: First, an initial online survey was employed to gather conventional retrospective measures, that is ex-post estimations of the participants on their media use characteristics. Subsequently, participants took part in a two week MESM study in order to measure media use in-situ. Strictly speaking, this sequencing makes a direct comparison between the two study parts problematic as the estimates do not refer to the same time frame. This approach was necessary, however, to avoid panel effects due to the reactivity of the design. Existing research supports that repeated self-observations in an ESM study increase self-awareness regarding the examined behavior and can influence later reports (e.g., CSIKSZENTMIHALYI/LARSON 1987; VAN DER ZOUWEN/VAN TILBURG 2001). The effect is expected to be small from one in-situ report to another, because people focus on specific instances of a behavior in each in-situ report (CARP/CARP 2007). Yet, the effect of repeated in-situ self-reports on a follow-up retrospective survey should be more detrimental. Indeed, in our extensive pretest we had found that participating in the MESM study prior to the online survey influenced respondents' answers. Apparently, they had learned to appraise their media use characteristics more correctly due to the ›self-observation period‹ they had undergone. We were less concerned about consistency or assimilation effects (i.e., prior questions impacting following ones) because of the nature of the measured concepts (behaviors and cognitions rather than attitudes) and the time lag between survey and MESM that leads to a wear-off effect (TOURANGEAU/RIPS/RASINSKI 2000: 207). Furthermore, to reduce probable bias, which might limit the external validity of comparing ex-post data to in-situ data, we made sure that the period under investigation did not include any special events (and nothing surprising came up either). Additionally, it can be stated that, while characteristics of specific situations in people's daily lives are highly diverse, the overall composition of situations that constitute people's daily routines are generally quite stable over time; thus, comparing MESM data to the data of a previous online survey was considered acceptable.

We considered usage of three different types of social media platforms to grasp exemplars of the above-mentioned particularities of mobile social media behavior: a chat app, a ›traditional‹ social networking site (SNS), and a video sharing platform. The duration of a usage episode is not universal across all types of social media. Rather it can be assumed to vary systematically across these types of platforms, dependent on their predominant mode of communication. Concentrating on the continuum between bidirectional and unidirectional modes of communication, we distinguish these types as follows:

The first type of platform enabling predominantly bidirectional communication is often termed chat apps (e.g., BENTON 2014). Chat apps mainly serve to mediate interpersonal communication in the form of (potentially) asynchronous and (mostly) written communication. They are characterized by a frequent change in the roles of sender and receiver. Due to these constant changes, single usage episodes can be assumed to be rather short but occur at a high frequency. One of the most widely used apps of this type in Germany is WhatsApp (TIPPELT/KUPFERSCHMITT 2015).

The second type does not only provide the possibility of mediated interpersonal communication but also enables a more broadcast style of one-to-many communication, i.e., sending out information to many receivers, not necessarily requiring feedback from them. These platforms can provide quick changes between sender and receiver roles but not necessarily require doing so. Hence, single usage episodes might be very short and highly frequent but with a higher variation in length as compared to the first type. Most SNS can be subsumed under this type, with Facebook being the most prominent one in Germany (SCHRÖDER 2016).

Third, there are social media platforms providing content sent by a relatively smaller number of senders to a broad audience of receivers, mostly video sharing platforms with YouTube being the most prominent example in Germany (KUPFERSCHMITT 2016). Due to the more broadcast-style nature of these platforms, length of usage episodes is probably higher and frequency of usage episodes is presumably lower than with the other types.

Thus, the study considers usage of three different types of social media platforms to grasp exemplars of the above-mentioned particularities of social media behavior, namely WhatsApp, Facebook, and YouTube.

5.2 *Participant sampling*

The sample consists of 126 students from three German universities. Participants consented in writing after being informed about the aims and the procedure of the study and received course credits as an incentive. A prerequisite of taking part in our study was ownership of a smartphone. Participants who did not complete the retrospective survey ($n = 7$) or who completed less than seven of the ESFs ($n = 7$) were excluded, leaving 112 participants (75.0% female, $M_{age} = 20.07$; $SD_{age} = 1.89$) for the analyses.

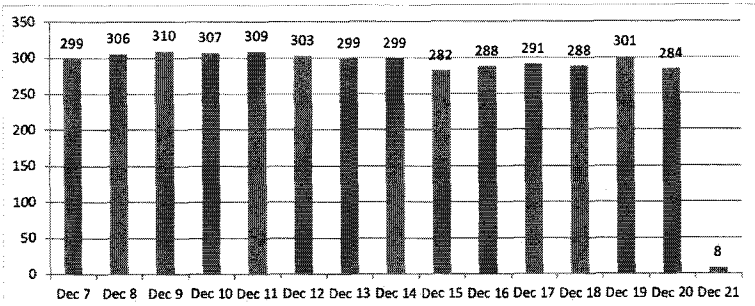
5.3 MESM *procedure and observation sampling*

The MESM-study lasted over a period of 14 days in December 2016. During this time each participant received three text messages per day randomly timed between 8.00 am and 10.00 pm. Each text message contained a link to a short online questionnaire (ESF) directly accessible via the participants' mobile Internet connection. The participants were asked to answer the ESF as soon as possible after being notified.

In total, we received 4246 completed ESFs. On average, participants completed 37.91 (SD = 6.56) of the 42 ESFs. The report latency between signal and actual participation time was 54.98 minutes (SD = 264.62, Mdn = 13.49). The ESFs were filled in 4.44 minutes on average (SD = 155.21, Mdn = 0.79). We excluded 70 ESFs with a completion duration of more than 10 minutes. Additionally, two ESFs were removed because two participants had stated never using Facebook in the retrospective survey and not completed the related questions but reported to use Facebook in one ESF per person. This procedure left 4174 ESFs with an average report latency of 55.09 minutes (SD = 266.51) and a completion duration of 0.97 minutes (SD = 0.84).

Overall participation in the study was satisfactory with a stable share of completed ESFs per day (see Figure 1).

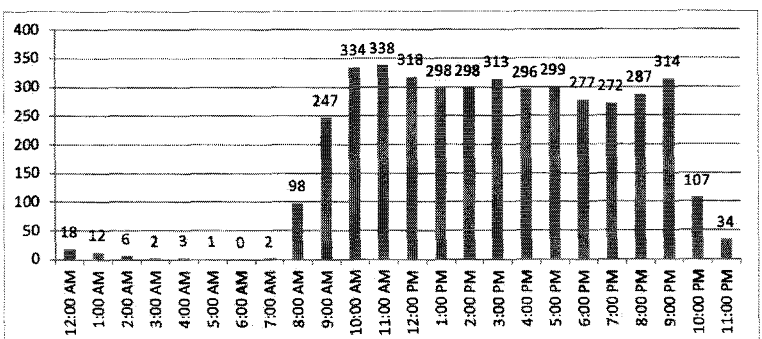
FIGURE 1
Completed ESFs over time by date



Note. N = 4174. Some respondents answered the last prompt of December 20 after midnight the next day (i.e., on December 21).

Regarding the spread of completed ESFs across the day, we observed a satisfactory distribution, showing a slight dent in the first hour. As no text message prompts were sent out between 10 pm and 8 am, there are significantly fewer completed ESFs in this period of time (see Figure 2).

FIGURE 2
Completed ESFs by hour of the day



Note. N = 4174. No prompts were sent out between 10:00 pm and 8:00 am.

The ESF determined whether or not YouTube, WhatsApp, and Facebook had been used in the last hour previous to answering the ESF. YouTube was used in 9.6 percent (n = 402), WhatsApp in 49.5 percent (n = 2068), and

Facebook in 18.0 percent ($n=751$) of the ESFs. Multiple choices were possible. Subsequently, participants were randomly assigned to one of the platforms they had stated to having used in the past hour and asked questions on their usage: 5.5 percent ($n=231$) of the final ESFs related to YouTube stemming from 67 participants, 40.1 percent ($n=1672$) to WhatsApp (from 110 participants), and 9.9 percent ($n=413$) related to Facebook (also from 100 participants). In 44.5 percent ($n=1858$) of the ESFs the participants did not use any of the three platforms.

5.4 *Measures*

5.4.1 *Measures in the retrospective survey*

An initial filter question recorded whether the respondents used the three media platforms. YouTube was used by all 112 respondents at least rarely. Two respondents stated never to use WhatsApp, four stated never to use Facebook. These participants were not asked on these respective platforms and are excluded from the analyses relating to the platforms. Respondents estimated the *duration of a regular usage episode* of YouTube, WhatsApp, and Facebook in minutes. The questionnaire specified that this should include how long respondents were occupied with the platform (by watching, reading, writing, posting) until they interrupted or ended the usage and turned towards a different activity.

Respondents stated their *usage frequency* of each platform on a scale ranging from (5) = ›several times per hour‹, (4) = ›several times per day‹, (3) = ›daily‹, (2) = ›at least once per week‹, (1) = ›more rarely‹ to (0) = ›never‹. Usage frequency was not distributed normally for the three platforms. Thus, we dichotomized the variables. 60.7 percent of the respondents used YouTube less than daily. 22.7 percent of the respondents used WhatsApp no more than several times per day. 26.9 percent of the respondents used Facebook no more than once a day.

On a bipolar five-point scale respondents rated whether they used the media platforms in *varying or stable circumstances*, referring to place (i.e., (1) ›always in the same place‹ to (5) ›always in different places‹), time, and social company. We dichotomized the variable (1) indicating that the platform was (mostly) used in a stable place, at the same time, and in a stable social company (values [4] and [5] of the original scale; for the measure of context stability NAAB/SCHNAUBER 2016). 75.9 percent of the respondents stated to use YouTube (mostly) at the same place, 30.4 percent (mostly) at the same time, and 73.2 percent (mostly) with the same social company. 3.6 percent of the respondents stated to use WhatsApp (mostly) at the same place, 0.9 percent (mostly) at the same time, and 8.2 percent (mostly) with the same social company. 20.4 percent of the respondents stated to use Facebook (mostly) at the same place, 10.2 percent (mostly) at the same time, and 49.1 percent (mostly) with the same social company.

TABLE 1

Descriptive statistics for the platform related measures in the retrospective survey

Measures	YouTube (*N* = 112)			WhatsApp (*N* = 110)			Facebook (*N* = 108)		
	M	*SD*	*α*	*M*	*SD*	*α*	*M*	*SD*	*α*
Duration of an average usage episode[1]	30.51	27.82	-	8.12	25.40	-	9.39	8.76	-
Usage frequency[2]	2.36	1.10	-	4.77	0.41	-	3.80	0.75	-
Habit strength[3]	2.70	0.80	.89	3.88	0.59	.81	3.21	0.76	.87
Perceived stability of usage place[4]	2.06	1.09	-	4.65	0.72	-	3.69	1.23	-
Perceived stability of usage time[5]	3.57	1.39	-	4.66	0.61	-	4.12	1.01	-
Perceived stability of social usage context[6]	1.96	1.14	-	4.08	1.05	-	2.72	1.32	-
involvement[7]	66.47	13.25	.87	71.82	10.94	.82	51.96	14.56	.88
Perceived social desirability of usage[8]	2.56	1.38	.85	2.55	1.22	.78	2.24	1.11	.65

Note.
N = 112 participants
1 in minutes
2 0 = never to 5 = several times per hour
3 1 = not at all habitual to 5 = fully habitual
4 1 = always at the same place; 5 = not always at the same place
5 1 = always at the same time; 5 = not always at the same time
6 1 = always with the same persons or alone; 5 = not always with the same persons or alone
7 0 = no involvement to 100 = high involvement
8 0 = not at all desirable to 4 = fully desirable

Habit strength of selecting each platform was measured with the self-report habit index by Verplanken and Orbell (2003) ranging from (1) = ›fully disagree‹ to (5) = ›fully agree‹ (12 items, e.g., ›I switch on [platform] automatically‹, ›Using [platform] belongs to daily routine‹, ›I start using [platform] before I realize I'm doing it‹).

Involvement in each platform was measured with a semantic differential scale containing nine items by Zaichkowsky (1994). The respondents used a slider from (0) = ›no involvement‹ to (100) = ›high involvement‹ (e.g., ›unimportant – important‹; ›irrelevant – relevant‹). *Perceived social desirability* of using the media platforms was measured by two items ranging from (0) = ›fully disagree‹ to (4) = ›fully agree‹ (›I sometimes have a bad conscious when using [platform]‹, ›I sometimes think I should not spend that much time using [platform]‹). We computed a mean score of the two items and reversed it, so that high values represent that using the platform is perceived as desirable.

The *tendency to give socially desirable answers* was measured by the social desirability scale by Stoeber (2001; $M = 0.62$; SD = 0.15; scale from 0 = none of 16 items that imply tendency to socially desirable answers answered affirmatively; 1 = all 16 items answered affirmatively).

Additionally, *age and gender* were measured.

The descriptives of the platform-related measures are specified in Table 1.

5.4.2 *Measures in the MESM-study*

In each ESF, the participants reported on the *duration of the last use episode* in minutes of the platform which they were assigned to in the respective ESF (YouTube: N = 67; $M = 26.60$; SD = 19.28; WhatsApp: N = 110; $M = 4.94$; SD = 3.93; Facebook: N = 100; $M = 7.98$; SD = 6.40).

A time stamp of the time of the prompt and of the actual participation time in each ESF were automatically saved to compute the *report latency*.

5.5 *Pretest*

In preparation of the study, we conducted an extensive pretest with two notifications per person per day over 14 days ($N = 71$ students; 86% female; $M_{age} = 22.2$, $SD_{age} = 2.9$). In total we received 1715 completed ESFs. There was a slight decline of completed ESFs over time and a slight dent in the morning hours. Overall participation in the study remained satisfactory, however. After the study was finished, the participants handed in feedback on the method and procedure. The procedure and the measures of the main study were optimized accordingly.

6. Results

6.1 *Preliminary analysis*

We computed the differences between the measures of media use duration obtained from the retrospective survey and from the in-situ mobile experience sampling method for the three platforms. (For an elaborate discussion of the differences of MSM use characteristics as a function of method of data collection and for detailed consideration of analyzing differences between the methods see NAAB/KARNOWSKI/SCHLÜTZ 2018.) For this purpose, the repeated in-situ observations of duration of usage episodes of the respondents obtained from the ESFs were aggregated to the level of an individual user. We used the arithmetic mean as an aggregation procedure. Then, we computed the difference between each individual's retrospective estimation of the duration of a usage episode and their aggregated in-situ estimations. The mean values of all three computed difference variables are positive, thus indicating that on average the respondents reported longer durations in retrospect compared to their averaged in-situ reports. The difference between the measures is greatest for YouTube ($N=67$; $M=10.39$; $SD=26.36$), followed by WhatsApp ($N=110$; $M=3.18$; $SD=25.47$), while it is smallest for Facebook ($N=100$; $M=1.40$; $SD=8.91$).

The fact that on average respondents estimated their media use duration in retrospect to be longer does not imply that all respondents showed such a pattern, though: For YouTube, 66 percent of the respondents estimated longer use durations in retrospect, while 29 percent estimated shorter durations in the retrospective survey than in the aggregated in-situ measures. For WhatsApp, 45 percent estimated longer durations in retrospect while 52 percent estimated a shorter duration. For Facebook, 53 percent estimated longer durations in retrospect and 37 percent shorter. These differences indicate the absolute level of inconsistency in estimation of a respondent's individual average usage duration. However, this inconsistency is relative to the length of the average duration (e.g., a difference of a minute between the retrospective estimation and the aggregated in-situ measure is of greater relevance when the average duration takes two minutes compared to when it takes 20 minutes). Thus, we standardized the difference between the retrospective value and the aggregated in-situ value to each individual's aggregated in-situ value. This value reflects the individual proportion of deviance in estimations. This deviance is 0.98 ($SD = 2.27$) for YouTube, 1.17 ($SD = 7.06$) for WhatsApp, and 0.66 ($SD = 1.80$) for Facebook. Here, too, we see an overall pattern of reporting higher values in retrospect: On average the difference between the retrospective value and the aggregated in-situ value is a little less than +100 percent of the aggregated in-situ duration for YouTube and about +120 percent for WhatsApp. Facebook users add about +66 percent of the duration when estimating the duration in retrospect.

The hypotheses outlined above propose influences of several indicators on the difference between retrospective and aggregated in-situ values of the duration of a usage episode. In line with theory the hypotheses suggest an influence of usage frequency (H1), habit strength (H2), context stability (H3), involvement (H4), and social desirability (H5) on duration of use. However, we do not postulate whether the predictor variables lead to differences resulting from higher retrospective or higher aggregated in-situ values. Thus, to test the hypotheses, we will refer to the absolute values of the proportional differences between the retrospective and the aggregated in-situ values. The absolute value of the proportional difference for YouTube is 1.29 ($SD = 2.11$), for WhatsApp 1.62 ($SD = 6.97$), and for Facebook 0.98 ($SD = 1.64$).[1] Thus, when we regard the relative inconsistencies without considering the direction, the inconsistencies are highest for WhatsApp, followed by YouTube, while they are lowest for Facebook.

6.2 *Explaining the differences*

To test the hypotheses, we calculated three OLS multiple regressions to predict the absolute values of the proportional difference between the

1 The means of the absolute values of the proportional difference differ from the proportion of inconsistencies stated above, because the absolute values have been computed on the individual level and then aggregated to a mean over all respondents.

retrospective and the aggregated in-situ value of the duration of a usage episode of each platform based on usage frequency, context stability, habit strength, involvement, and the interaction between tendency to social desirable answers and perceived social desirability of each platform.

TABLE 2

Hierarchical multiple regression analyses

predicting the absolute values of the proportional difference between the retrospective and the aggregated in-situ value of duration of a usage episode for YouTube, WhatsApp, and Facebook from usage frequency, habit strength, context stability, involvement, and social desirability

	Absolute values of the proportional difference between the retrospective and the aggregated in-situ value of the duration of a usage episode					
	YouTube		WhatsApp		Facebook	
Predictor	ΔR^2	ß	ΔR^2	ß	ΔR^2	ß
Step 1	.008		.010		.016	
Number of ESFs of the respondent		.023		.085		-.083
Report latency		.167		.083		-.107
Step 2	.238		.075		.085	
Usage frequency		-.047		0.063		-.150
Habit strength		.097		.133		.080
Stability of place		.060		.049		.095
Stability of time		.062		.031		.007
Stability of social company		.331		-.056		.084
Involvement		-.350*		.041		-.123
Tendency to social desirable answers		.142		-.355		-.092
Perceived social desirability of platform		.179		-.615		-.386
Interaction: tendency to social desirable answers * perceived social desirability		.139		.646		.516
Total R^2	.246		.085		.101	
(corr. R^2)	(.095)		(-.018)		(-.011)	
N	67		110		100	

Note.
* $p<.05$; ** $p<.01$; *** $p<.001$.

In all regressions, we controlled for the number of filled ESFs per respondent ($M=37.91$; $SD=6.56$, see above) out of the maximum number of 42 prompted ESFs as an indicator of compliance with the study. Additionally, we controlled for the average report latency for the ESFs of the platform in question. That is, we computed the arithmetic mean of the time spans between the prompt and the actual participation in an ESF over all ESFs of a respondent referring to YouTube, WhatsApp, or Facebook. Since the report latencies for all platforms were extremely skewed to the right, we dichotomized the variables (YouTube: 52.2% of the respondents answered within 15 min after the prompt on average; WhatsApp: 32.7%; Facebook: 59.0%). Aggregated report latency was included in the regression predicting the difference between the retrospective and the aggregated in-situ values of the usage duration. We inserted the control variables as a first step, followed by a second step including all predictor variables.

Overall, the regressions were not significant, neither for YouTube ($F(11,55)=1.629$, $p=.116$; $R^2=.246$), WhatsApp ($F(11,98)=0.825$, $p=.615$; $R^2=0.085$), nor Facebook ($F(11,88)=0.902$, $p=.542$; $R^2=0.101$) (see Table 2).[2]

7. Discussion

This study was concerned with the quantification of media use as a methodological challenge, notably with regard to mobile social media. The relevance of accurately measuring media use is based on the importance of the variable in many areas of communication research (DE VREESE/NEIJENS 2016; SLATER 2004). Thus, if this parameter is not quantified accurately, associations with or dependencies on other concepts might be misjudged. Standardized retrospective surveys that are commonly employed to capture media use are prone to measurement error. Errors might occur on several steps of the cognitive process from survey question to response (SCHWARZ/OYSERMAN/PEYTCHEVA 2010). Particularly recalling relevant behavior and estimating averages of behavior in retrospect are prone to mistakes. The longer ago or the less remarkable the period of time on which the partici-

2 To control for effects of standardization on our results, we also ran our analyses using the non-standardized differences between the retrospective and the aggregated in-situ values as dependent variables. These regressions yielded no significant results for all three platforms either.

pants are asked to report, the more difficult it is for them to give an accurate answer. This is especially true with regard to the various kinds of mobile social media use that are characterized by ubiquitous access to a wide variety of platforms in all kinds of situations and assembling a multitude of inter- and intraindividually different usage patterns.

The leading rationale of this study was that in-situ self-reports demand less cognitive abilities during retrieval. Additionally, repeated in-situ self-reports overcome the challenge that single reports on recent media behavior might be biased by the specifics of the situation. Thus, we expected that repeated in-situ measures resulted in valid measures of mobile social media use. To test differences between media use measures obtained from retrospective survey and MESM, respectively, we compared retrospective survey data and MESM data on the use of three different types of social media platforms, namely YouTube, WhatsApp, and Facebook. As elaborated by Naab, Karnowski, and Schlütz (2018) the estimates of usage duration differed systematically: The values in the retrospective survey are higher compared to the aggregated in-situ measurement. The relative differences between retrospective survey and aggregated in-situ measurements are highest for WhatsApp, followed by YouTube, and lowest for Facebook. The study did not provide us with a benchmark which method of data collection is more valid and produces ›true‹ scores. However, from the theoretical assumptions we might deduce that the MESM values are more trustworthy. Yet, given the outlined deficits of MESM – among them being a biased sample as well as potential reactivity effects in the longitudinal design – this should be considered with care. Additionally, the facts that not all ESFs had been answered and that there was considerable report latency between the prompts and answering the ESFs indicate weaknesses of the MESM results. The data show that neither compliance rate nor report latency influenced the differences between retrospective and MESM data. Still, validating the retrospective and MESM data with data from tracking participants' use of certain mobile social media applications seems a valuable approach for future research.

We tried to identify factors that might influence the difference between retrospective and aggregated in-situ duration values. From the cognitive process of answering survey questions we deduced several characteristics of media behavior which were assumed to influence these differences: frequency of use, habit strength, context stability, involvement, and tendency to social desirable answers. But none of the independent variables could explain the differences we found between retrospective and aggregated in-situ duration estimates. For one, this might be due to measurement issues with the assumed predictor variables. These were measured in the retrospective survey. While this is inevitable for the respondents' tendency to social desirability, which is not supposed to vary across situations, the further factors context stability, habit strength, and involvement might vary intra-individually. Frequency of media use cannot be measured in-situ, but still it might be difficult to estimate for the respondents. Thus, most of the presumed predictor variables might be prone to the same measurement errors as discussed above. This could reduce their explanatory power. Future studies will have to examine whether the pattern of deviance found here for the duration of a usage episode also holds for further constructs like, for instance, gratifications.

Apart from this methodological explanation, the presumed factors might just be irrelevant for the discrepancies between the methods. This raises the question whether answering behavior is to some extent random or whether other factors can explain the differences observed. Literature on how respondents answer self-report questions has barely considered individuals' heuristics employed to recall and estimate use duration (see for exceptions CONRAD/BROWN/CASHMAN 1998; MENON 1993; SCHWARZ 2007). From the data at hand we might speculate that respondents do not consider all instances of a mobile social media use equally when arriving at estimations about their usual length of a usage episode. Probably episodes from just a little while ago are given more weight than earlier ones. Also usage episodes with particularly notable media content could dominate during estimation. In this vein, Potter (2008) proposes four exposure states (attentional, automatic, transported, and self-reflexive) as a feature of usage episodes that influence how people experience and, even more important to this study, how they remember media use.

Additionally, answering strategies may vary inter-individually. In this regard, Schwarz and Oyserman (2001) refer to variations in the respondent's processing strategies. That means different people might employ different strategies to arrive at retrospective estimations about their usual or typical behavior. Individuals might have a general tendency to optimize or satisfice, respectively, when answering self-report questions. Future studies should examine the influence of personality characteristics (apart from the here tested tendency to social desirability) on estimation strategies and explore whether these affect discrepancies between retrospective and aggregated in-situ values. In this regard, future studies also need to validate the results with representative samples. Respondents with more heterogeneous use of mobile social media and varying cognitive strategies and abilities might compute estimations differently.

Finally, future studies need to test whether the non-influence of our presumed predictor variables also holds true for traditional and for other mobile social media. This seems a relevant endeavor since Naab, Karnowski, and Schlütz (2018) find a fairly systematic pattern of deviance between retrospective and in-situ values for several further mobile social media use characteristics beyond duration of a usage episode.

References

BAYER, J. B.; ELLISON, N. B.; SCHOENEBECK, S. Y.; FALK, E. B.: Sharing the small moments: Ephemeral social interaction on Snapchat. In: *Information, Communication & Society*, *19*(7), 2015, pp. 956-977. https://doi.org/10.1080/1369118X.2015.1084349

BENTON, J.: *Here's some remarkable new data on the power of chat apps like WhatsApp for sharing news stories* (2014, November 20). http://www.niemanlab.org/2014/11/heres-some-remarkable-new-data-on-the-power-of-chat-apps-like-whatsapp-for-sharing-news-stories/

BERNARD, R. H.; KILLWORTH, P.; KRONENFELD, D.; SAILER, L.: The problem of informant accuracy: The validity of retrospective data. In: *Annual Review of Anthropology*, *13*(1), 1984, pp. 495-517. https://doi.org/10.1146/annurev.anthro.13.1.495

BOASE, J.; KOBAYASHI, T.: Mobile communication networks in Japan and America. In: *China Media Research*, *8*, 2012, pp. 90-98

BOASE, J.; LING, R.: Measuring mobile phone use: Self-report versus log data. In: *Journal of Computer-Mediated Communication*, *18*(4), 2013, pp. 508-519. https://doi.org/10.1111/jcc4.12021

BOLGER, N.; DAVIS, A.; RAFAELI, E.: Diary methods: capturing life as it is lived. In: *Annual Review of Psychology*, *54*, 2003, pp. 579-616. https://doi.org/10.1146/annurev.psych.54.101601.145030

BOUWMAN, H.; DE REUVER, M.; HEERSCHAP, N.; VERKASALO, H.: Opportunities and problems with automated data collection via smartphones. In: *Mobile Media & Communication*, *1*, 2013, pp. 63-68. doi: 10.1177/2050157912464492

BRADBURN, N.; RIPS, L.; SHEVELL, S.: Answering autobiographical questions: The impact of memory and inference on surveys. In: *Science, 236*(4798), 1987, pp. 157-161. https://doi.org/10.1126/science.3563494

BRANDT, J.; WEISS, N.; KLEMMER, S. R.: Lowering the burden for diary studies under mobile conditions. In: ACM (Ed.): *CHI'07 extended abstracts on human factors in computing systems*. San Jose [ACM] 2007, pp. 2303-2308

CARP, F. M.; CARP, A.: The validity, reliability and generalizability of diary data. In: *Experimental Aging Research, 7*(3), 2007, pp. 281-296. https://doi.org/10.1080/03610738108259811

CHANG, L.; KROSNICK, J. A.: Measuring the frequency of regular behaviors: Comparing the »typical week« to the »past week«. In: *Sociological Methodology, 33*(1), 2003, pp. 55-80. https://doi.org/10.1111/j.0081-1750.2003.t01-1-00127.x

COHEN, A. A.; LEMISH, D.: Real time and recall measures of mobile phone use: Some methodological concerns and empirical applications. In: *New Media & Society, 5*(2), 2003, pp. 167-183. https://doi.org/10.1177/1461444803005002002

COHEN, E. L.; BOWMAN, N. D.; LANCASTER, A. L.: R U with Some1? Using text message experience sampling to examine television coviewing as a moderator of emotional contagion effects on enjoyment. In: *Mass Communication and Society, 19*(2), 2015, pp. 149-172. https://doi.org/10.1080/15205436.2015.1071400

CONNER, T.; BLISS-MOREAU, E.: Sampling human experience in naturalistic settings. In: HESSE-BIBER, S.; LEAVY, P. (Eds.): *Emergent methods in social research*. Thousand Oaks, CA [Sage] 2006, pp. 109-129

CONRAD, F. G.; BROWN, N. R.; CASHMAN, E. R.: Strategies for estimating behavioral frequency in survey interviews. In: *Memory, 6*(4), 1998, pp. 339-366. doi: 10.1080/741942603

CSIKSZENTMIHALYI, M.; LARSON, R.: Validity and reliability of the experience-sampling method. In: *The Journal of Nervous and Mental Disease, 175*(9), 1987, pp. 526-536. https://doi.org/10.1097/00005053-198709000-00004

CROWNE, D. P.; MARLOWE, D.: A new scale of social desirability independent of psychopathology. In: *Journal of Consulting Psychology, 24*(4), 1960, pp. 349-354. http://dx.doi.org/10.1037/h0047358

DANNER, U. N.; AARTS, H.; VRIES, N. K. DE: Habit vs. intention in the prediction of future behaviour: the role of frequency, context stability and mental accessibility of past behaviour. In: *The British Journal of Social Psychology, 47*(Pt 2), 2008, pp. 245-265. https://doi.org/10.1348/014466607X230876

DEMAIO, T. J.: Social desirability and survey measurement: A review. In: TURNER, C. F.; MARTIN, E. (Eds.): *Surveying subjective phenomenon*. New York, NY [Russell Sage] 1985, pp. 257-282

DE VREESE, C. H.; NEIJENS, P.: Measuring media exposure in a changing communications environment. In: *Communication Methods and Measures, 10*(2-3), 2016, pp. 69-80. https://doi.org/10.1080/19312458.2016.1150441

FELDMANN, V.: *Leveraging mobile media: Cross-media strategy and innovation policy for mobile media communication*. Heidelberg [Physica-Verlag] 2005

FISHBEIN, M.; HORNIK, R.: Measuring media exposure: An introduction to the special issue. In: *Communication Methods and Measures, 2*, 1-5, 2008. doi: 10.1080/19312450802095943

HA, L.; HU, X.; FANG, L.; HENIZE, S.; PARK, S.; STANA, A.; ZHANG, X.: Use of survey research in top mass communication journals 2001-2010 and the total survey error paradigm. In: *Review of Communication, 15*(1), 2015, pp. 39-59. https://doi.org/10.1080/15358593.2015.1014401

HAMPTON, K. N.; LIVIO, O.; SESSIONS GOULET, L.: The social life of wireless urban spaces: Internet use, social networks, and the public realm. In: *Journal of Communication, 60*(4), 2010, pp. 701-722. https://doi.org/10.1111/j.1460-2466.2010.01510.x

HARGITTAI, E.; HINNANT, A.: Digital inequality: Differences in young adults' use of the Internet. In: *Communication Research, 35*(5), 2008, pp. 602-621. https://doi.org/10.1177/0093650208321782

HESS, T.: Intermedia – Struktureller Wandel der Intermediation in der Medienbranche [Intermedia – Structural change of intermediation in the media industry]. In: HESS, T. (Ed.): *Ubiquität, Interaktivität, Konvergenz und die Medienbranche* [Ubiquity, interactivity, convergence and the media industry]. Göttingen [Univ.-Verl. Göttingen] 2007, pp. 109-126

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

KAHLOR, L.; DUNWOODY, S.; GRIFFIN, R. J.; NEUWIRTH, K.; GIESE, J.: Studying heuristic-systematic processing of risk communication. In: *Risk Analysis*, *23*(2), 2003, pp. 355-368. https://doi.org/10.1111/1539-6924.00314

KARNOWSKI, V.; JANDURA, O.: When lifestyle becomes behavior: A closer look at the situational context of mobile communication. In: *Telematics and Informatics*, *31*(2), 2014, pp. 184-193. https://doi.org/10.1016/j.tele.2013.11.001

KARNOWSKI, V.; KÜMPEL, A. S.; LEONHARD, L.; LEINER, D. J.: From incidental news exposure to news engagement. How perceptions of the news post and news usage patterns influence engagement with news articles encountered on Facebook. In: *Computers in Human Behavior*, *76*, 2017, pp. 42-50. https://doi.org/10.1016/j.chb.2017.06.041

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

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

KUBEY, R.; CSIKSZENTMIHALYI, M. (Eds.): *Television and the quality of life*. Hillsdale, NJ [Lawrence Erlbaum] 1990

KUPFERSCHMITT, T.: Online-Videoreichweite steigt bei weiter geringer Nutzungsdauer [Online video reach increases at still low usage duration]. In: *Media Perspektiven*, 9/2016, pp. 448-459

LAROSE, R.: The problem of media habits. In: *Communication Theory*, *20*(2), 2010, pp. 194-222. https://doi.org/10.1111/j.1468-2885.2010.01360.x

LARSON, R.; CSIKSZENTMIHALYI, M.: The experience sampling method. In: REIS, H. T. (Ed.): *Naturalistic approaches to studying social interaction*. San Francisco, CA [Jossey-Bass.] 1983, pp. 41-56

LEE, C.-J.; HORNIK, R.; HENNESSY, M.: The reliability and stability of general media exposure measures. In: *Communication Methods and Measures*, *2*(1-2), 2008, pp. 6-22. https://doi.org/10.1080/19312450802063024

LEE, E.; HU, M. Y.; TOH, R. S.: Are consumer survey results distorted? Systematic impact of behavioral frequency and duration on survey

response errors. In: *Journal of Marketing Research, 37*(1), 2000, pp. 125-133. https://doi.org/10.1509/jmkr.37.1.125.18724
LING, R.: *Taken for grantedness: The embedding of mobile communication into society*. Cambridge, MA [MIT Press] 2012
MENON, G.: The effects of accessibility on information in memory on judgements of behavioral frequencies. In: *Journal of Consumer Research, 20*(3), 1993, pp. 431-440
NAAB, T. K.; SCHNAUBER, A.: Habitual initiation of media use and a response-frequency measure for its examination. In: *Media Psychology, 19*(1), 2016, pp. 126-155. https://doi.org/10.1080/15213269.2014.951055
NAAB, T. K.; KARNOWSKI, V.; SCHLÜTZ, D.: Reporting Mobile Social Media Use: How Survey and Experience Sampling Measures Differ. In: *Communication Methods and Measures*, 2018. https://doi.org/10.1080/19312458.2018.1555799
NIEDERDEPPE, J.: Meeting the challenge of measuring communication exposure in the digital age. In: *Communication Methods and Measures, 10*(2-3), 2016, pp. 170-172. https://doi.org/10.1080/19312458.2016.1150970
NISBETT, R. E.; WILSON, T. D.: Telling more than we can know: Verbal reports on mental processes. In: *Psychological Review, 84*(3), 1977, pp. 231-259. https://doi.org/10.1037//0033-295x.84.3.231
PALEN, L.; SALZMAN, M.: Voice-mail diary studies for naturalistic data capture under mobile conditions. In: ACM (Ed.): *Proceedings of the 2002 ACM conference on Computer supported cooperative work*. New Orleans, LA [ACM] 2002, pp. 87-95
PAULHUS, D. L.; HARMS, P. D.; BRUCE, M. N.; LYSY, D. C.: The over-claiming technique: Measuring self-enhancement independent of ability. In: *Journal of Personality and Social Psychology, 84*(4), 2003, pp. 890-904 https://doi.org/10.1037/0022-3514.84.4.890
POTTER, W. J.: The importance of considering exposure states when designing survey research studies. In: *Communication Methods and Measures, 2*(1-2), 2008, pp. 152-166. doi: 10.1080/19312450802062299
PRIOR, M.: Improving media effects research through better measurement of news exposure. In: *The Journal of Politics, 71*(3), 2009a, pp. 893-908. https://doi.org/10.1017/S0022381609090781
PRIOR, M.: The immensely inflated news audience: Assessing bias in self-reported news exposure. In: *Public Opinion Quarterly, 73*(1), 2009b, pp. 130-143. https://doi.org/10.1093/poq/nfp002

QUANDT, T.; VON PAPE, T.: Living in the mediatope: A multimethod study on the evolution of media technologies in the domestic environment. In: *The Information Society, 26*(5), 2010, pp. 330-345. https://doi.org/10.1080/01972243.2010.511557

REIS, H. T.; GABLE, S. L.: Event-sampling and other methods for studying everyday experience. In: REIS, H. T.; JUDD, C. M. (Eds.): *Handbook of research methods in social and personality psychology*. New York, NY [Cambridge University Press] 2000, pp. 190-222

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

ROSA, H.: *Social acceleration. A new theory of modernity*. Ney York, NY [Columbia University Press] 2015

ROSS, M.: Relation of implicit theories to the construction of personal histories. In: *Psychological Review, 96*(2), 1989, pp. 341-357. https://doi.org/10.1037//0033-295X.96.2.341

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

SCHERER, H.; NAAB, T. K.: Messen im Zeitverlauf [Measuring in the course of time]. In: MÖHRING, W.; SCHLÜTZ, D. (Eds.): *Handbuch standardisierte Erhebungsmethoden in der Kommunikationswissenschaft* [Handbook of standardized methods in communication science]. Wiesbaden [VS] 2013, pp. 103-123

SCHERPENZEEL, A. C.; SARIS, W. E.: The validity and reliability of survey questions: A meta-analysis of MTMM Studies. In: *Sociological Methods & Research, 25*(3), 1997, pp. 341-383. https://doi.org/10.1177/0049124197025003004

SCHLÜTZ, D.: *Bildschirmspiele und ihre Faszination: Zuwendungsmotive, Gratifikationen und Erleben interaktiver Medienangebote* [Video games and their fascination: Motives, gratifications, and experience of interactive media]. München [Reinhard Fischer] 2002

SCHLÜTZ, D.; SCHERER, H.: Der Einsatz der ›Experience Sampling Method‹ in der Medienwissenschaft [Using the ›Experience sampling method‹ in media research]. In: *Zeitschrift für Medienpsychologie, 13*(3), 2001, pp. 146-149. https://doi.org/10.1026//1617-6383.13.3.146

SCHNAUBER, A.: *Medienselektion im Alltag: Die Rolle von Gewohnheiten im Selektionsprozess [Media selection in everyday life: The role of habits in the selection process]*. Wiesbaden [Springer VS] 2017

SCHRÖDER, J.: *Die populärsten sozialen Netzwerke in Deutschland: Facebook klar vorn, Instagram, Snapchat und musical.ly boomen* [The most popular social networks in Germany: Facebook clearly at the top, Instagram, Snapchat and musical.ly boom] 2016. http://meedia.de/2016/12/01/die-20-populaersten-sozialen-netzwerke-in-deutschland-facebook-klar-vorn-instagram-snapchat-und-musical-ly-boomen/

SCHWARZ, N.: Retrospective and concurrent self-reports: The rationale for real-time data capture. In: STONE, A.; SHIFFMAN, S. S.; ATIENZA, A.; NEBELING, L. (Eds.): *The science of real-time data capture: Self-reports in health research*. New York, NY [Oxford University Press] 2007, pp. 11-26

SCHWARZ, N.; OYSERMAN, D.: Asking questions about behavior: Cognition, communication, and questionnaire construction. In: *American Journal of Evaluation*, *22*(2), 2001, pp. 127-160. https://doi.org/10.1177/109821400102200202

SCHWARZ, N.; OYSERMAN, D.; PEYTCHEVA, E.: Cognition, communication, and culture: Implications for the survey response process. In: HARKNESS, J. A.; BRAUN, M.; EDWARDS, B.; JOHNSON, T. P; LYBERG, L. E.; MOHLER, P. P.; SMITH, T. W. (Eds.): *Survey methods in multinational, multiregional and multicultural contexts*. Hoboken, NJ [John Wiley] 2010, pp. 177-190

SCOLLON, C. N.; KIM-PRIETO, C.: Experience sampling: Promises and pitfalls, strengths and weaknesses. In: *Journal of Happiness Studies*, *4*(1), 2003, pp. 5-34. https://doi.org/10.1023/a:1023605205115

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

STOEBER, J.: The Social Desirability Scale-17 (SDS-17): Convergent validity, discriminant validity, and relationship with age. In: *European Journal of Psychological Assessment*, *17*(3), 2001, pp. 222-232. doi:10.1027//1015-5759.17.3.222

STRUCKMANN, S.; KARNOWSKI, V.: News consumption in a changing media ecology: An MESM-study on mobile news. In: *Telematics and Informatics*, *33*(2), 2016, pp. 309-319. https://doi.org/10.1016/j.tele.2015.08.012

TIPPELT, F.; KUPFERSCHMITT, T.: Social Web: Ausdifferenzierung der Nutzung – Potenziale für Medienanbieter [Social web: Differentiation of usage – Potential for media providers]. In: *Media Perspektiven*, 10/2015, pp. 442-452

TOURANGEAU, R.; RIPS, L. J.; RASINSKI, K.: *The psychology of survey response*. Cambridge, MA [University Press] 2000

VALKENBURG, P. M.; PETER, J.: Five challenges for the future of media-effects research. In: *International Journal of Communication, 7*, 2013, pp. 197-215. doi: 1932-8036

VANDEN ABEELE, M.; BEULLENS, K.; ROE, K.: Measuring mobile phone use: Gender, age and real usage level in relation to the accuracy and validity of self-reported mobile phone use. In: *Mobile Media & Communication, 1*(2), 2013, pp. 213-236. https://doi.org/10.1177/2050157913477095

VAN DER ZOUWEN, J.; VAN TILBURG, T.: Reactivity in Panel Studies and its Consequences for Testing Causal Hypotheses. In: *Sociological Methods & Research, 30*(1), 2001, pp. 35 - 56. https://doi.org/10.1177/0049124101030001003

VERPLANKEN, B.; ORBELL, S.: Reflections on past behavior: A self-report index of habit strength. In: *Journal of Applied Social Psychology, 33*(6), 2003, pp. 1313-1330. https://doi.org/10.1111/j.1559-1816.2003.tb01951.x

VORDERER, P.: Der mediatisierte Lebenswandel [The mediatized lifestyle: Permanently online, permanently connected]. In: *Publizistik, 60*(3), 2015, pp. 259-276. https://doi.org/10.1007/s11616-015-0239-3

VORDERER, P.; KRÖMER, N.; SCHNEIDER, F.: Permanently online – permanently connected: Explorations into university students' use of social media and mobile smart devices. In: *Computers in Human Behavior, 63*, 2016, pp. 694-703. https://doi.org/10.1016/j.chb.2016.05.085

WESTLUND, O.: From mobile phone to mobile device: News consumption on the go. In: *Canadian Journal of Communication, 33*(3), 2008, pp. 443-463. https://doi.org/10.22230/cjc.2008v33n3a2004

WIRTH, W.: Involvement. In: BRYANT, J.; VORDERER, P. (Eds.): *Psychology of entertainment*. Mahwah, NJ [Lawrence Erlbaum] 2006, pp. 199-213

ZAICHKOWSKY, J. L.: The Personal Involvement Inventory. Reduction, revision, and application to advertising. In: *Journal of Advertising, 23*(4), 1994, pp. 59-70. https://doi.org/10.1080/00913367.1943.10673459