Check for updates

Special Issue Article

Measuring Smartphone Use: Survey Versus Digital Behavioral Data

Social Science Computer Review
2025, Vol. 43(5) 1030–1049
© The Author(s) 2024

Article reuse guidelines:
sagepub.com/journals-permissions
DOI: 10.1177/08944393231224540
journals.sagepub.com/home/ssc

Alexander Wenz[1], **Florian Keusch**[1], **and Ruben L. Bach**[1]

Abstract

While digital technology use and skills have typically been measured with surveys, digital behavioral data that are passively collected from individuals' digital devices have recently emerged as an alternative method of measuring technology usage patterns in a more unobtrusive and detailed way. In this paper, we evaluate how passively collected smartphone usage data compare to self-reported measures of smartphone use, considering the three usage dimensions *amount of use*, *variety of use*, and *activities of use*. Based on a sample of smartphone users in Germany who completed a survey and had a tracking app installed on their smartphone, we find that the alignment between the survey and digital behavioral data varies by dimension of smartphone use. Whereas *amount of use* is considerably overreported in the survey data, *variety of use* aligns more closely across the two data sources. For *activities of use*, the alignment differs by type of activity. The results also show that the alignment between survey and digital behavioral data is systematically related to individuals' sociodemographic characteristics, including age, gender, and educational attainment. Finally, latent class analyses conducted separately for the survey and digital behavioral data suggest similar typologies of smartphone use, although the overlap between the typologies on the individual level is rather small.

Keywords

digital trace data, smartphone tracking, survey data, self-report, latent class analysis, user typology

Introduction

Over the last decade, smartphones have become one of the most frequently used devices for accessing the Internet (Rice et al., 2023), but inequalities in the use of smartphone technologies and in smartphone-related skills persist in the general population (Wenz & Keusch, 2023). In research about the digital divide, the vast majority of studies have relied on surveys for measuring

[1]University of Mannheim, Germany

Corresponding Author:
Alexander Wenz, Mannheim Centre for European Social Research, University of Mannheim, A5, 6, 68131 Mannheim, Germany.
Email: a.wenz@uni-mannheim.de

Data Availability Statement included at the end of the article

digital technology use and skills (e.g., Büchi et al., 2016; Hargittai, 2005; van Deursen et al., 2016). Surveys are a relatively cost-effective approach for collecting data on population-representative samples and are still the predominant method of data collection in the social sciences (Sturgis & Luff, 2021). However, surveys also have limitations with regard to representation and measurement. For example, self-reports of behaviors, such as digital technology use, are prone to measurement error due to recall error, social desirability, and other prevalent errors in self-reports, potentially leading to biased results (Tourangeau et al., 2000).

Digital behavioral data (DBD) that are passively collected from individuals' digital devices have recently emerged as an alternative method for collecting data about attitudes and behaviors (Keusch & Kreuter, 2022). On smartphones, for example, the activities carried out on the device, including voice calls, text messages, and app usage, are continuously captured through smartphone usage logs by the operating system (Harari et al., 2016). Compared to surveys, DBD allow researchers to measure smartphone usage patterns not only unobtrusively, that is, not relying on the respondents' self-reports but also in a more detailed way and over a longer period of time.

In this paper, we evaluate how smartphone usage data that are passively collected from smartphones compare to self-reported measures of smartphone use based on a sample of smartphone users in Germany. First, we examine to what extent the DBD- and survey-based usage measures align with each other and how the alignment varies by sociodemographic characteristics. Second, we use latent class analysis to identify smartphone usage types separately for the DBD- and survey-based measures and compare the results across the two methods of data collection.

Background

In digital divide research, surveys have been the most frequently employed approach for measuring Internet and digital technology use. Three dimensions of use have typically been distinguished (Blank & Groselj, 2014): amount of use, variety of use, and activities of use. While *amount of use* has been measured with survey questions about the frequency of going online (e.g., number of hours spent online per day), the frequency of engaging in different online activities (e.g., number of hours spent on social media platforms per day), or total years of Internet use (e.g., van Deursen & van Dijk, 2014), *variety of use* has typically been measured as the number of different online activities that individuals report (e.g., Reisdorf & Groselj, 2017). *Activities of use*, in turn, have been measured with survey questions about which types of online activities individuals engage in, such as information seeking, communication, or entertainment (e.g., Zillien & Hargittai, 2009).

To conceptualize and classify the large variation in Internet and digital technology use, usage typologies have been constructed, typically based on latent class analysis or other types of cluster analyses (Blank & Groselj, 2014). In a study of German smartphone owners, for example, Wenz and Keusch (2023) identified six distinct types of users who vary strongly in their self-reported frequency of smartphone use, self-rated smartphone skills, and activities carried out on their smartphone: (1) *advanced users* (who use their smartphone frequently, for all examined activities, and have advanced skills), (2) *broad non-social media users* (who use their smartphone frequently, for a large variety of activities except for social media, and have intermediate or advanced skills), (3) *broad non-commercial users* (who use their smartphone frequently, for a large variety of activities except for online banking and purchases, and have beginner-level, intermediate, or advanced skills), (4) *social media and information users* (who use their smartphone frequently, mainly for browsing websites, reading or writing emails, and using social media, and have beginner-level skills), (5) *basic general users* (who use their smartphone frequently, for taking photos, browsing websites, using GPS/location-aware apps, and reading or writing emails, and have beginner-level or intermediate skills), and (6) *camera users* (who use their smartphone less

frequently, mainly for taking photos, and have beginner-level skills). The smartphone usage types also differed significantly by sociodemographic characteristics, with the types reflecting more frequent and diverse usage patterns being younger and having higher levels of educational attainment.

More recently, DBD that are collected as a by-product of individuals' everyday interactions with digital devices have emerged as an alternative approach for collecting data about Internet and digital technology use. While DBD have yet received little attention in digital divide research, they are increasingly used in related disciplines, such as psychology and communication research, for measuring Internet and smartphone use (e.g., Bach & Wenz, 2020; Ellis, 2019; Festic et al., 2021; Harari et al., 2020; Schoedel et al., 2020), social media use (e.g., Guess et al., 2019; Haenschen, 2020; Mahalingham et al., 2023), or news consumption (e.g., Barthel et al., 2020; Möller et al., 2020; Stier et al., 2022). Various techniques can be applied for collecting individual-level DBD on digital technology use (Ohme et al., 2023). Many studies follow a *tracking* approach in which participants are asked to download a research app on their mobile devices that passively captures prospective data about the apps used and the websites visited on the device (e.g., Araujo et al., 2017; Festic et al., 2021). Recent studies have also suggested a *data donation* approach in which participants are requested to donate their retrospective app usage data by manually recording usage reports that appear on the smartphone, such as the iOS Screen Time feature (e.g., Baumgartner et al., 2023; Jones-Jang et al., 2020; Ohme et al., 2021).

A small number of typologies of Internet and digital media use have been constructed based on DBD, mostly in the area of marketing research (Chen et al., 2019; Hamka et al., 2014; Lee et al., 2018). For example, Hamka et al. (2014) conducted a latent class analysis based on log files from 129 smartphone users who installed a tracking app on their device for at least two weeks. They identified six types of smartphone users who vary in the number of URLs visited, the number of apps used, and the number of new apps installed: (1) *application ignorant users* (who visit a small number of URLs in the mobile browser and use a small number of apps per day), (2) *basic application users* (who visit a small number of URLs but use a medium number of apps), (3) *average application users* (who visit a medium number of URLs and use a medium number of apps), (4) *information seekers* (who visit a larger number of URLs but use a small number of apps), (5) *app savvy users* (who visit a large number of URLs, use an extensive number of apps, and install a large number of new apps), and (6) *high utility users* (who visit an extensive number of URLs, use an extensive number of apps, but install a small number of new apps).

A growing body of research has examined to what extent DBD- and survey-based measures of Internet and digital technology use align with each other, mostly focusing on amount of use and treating DBD as the gold standard. Generally, the correlations between self-reported and tracked measures of digital technology use were found to be small to moderate (e.g., Andrews et al., 2015; Araujo et al., 2017; Boase & Ling, 2013; Deng et al., 2019; Ellis et al., 2019; Parry et al., 2021; Revilla et al., 2017; Scharkow, 2016). However, the findings about the direction of the difference, that is, whether individuals over- or underreport their digital technology use in surveys compared to their tracked behavior, are rather mixed.

In this paper, we study how DBD compare to survey data for measuring smartphone use and expand upon existing research by examining the three usage dimensions that are typically distinguished in digital divide research: amount of use, variety of use, and activities of use. We refrain from treating DBD as the gold standard since these data were also shown to have errors (Bosch & Revilla, 2022) and instead investigate the level of alignment between the survey and the DBD, relying on three indicators: absolute error (any differences in the measured behavior between the survey and the DBD), underreporting (behavior is observed in the DBD but not indicated in the survey), and overreporting (behavior is indicated in the survey but not observed in

the DBD). In addition, we investigate whether DBD and survey data collected from the same individuals lead to different user typologies.

Data and Methods

Sample

The DBD and survey data were collected as part of the project "Political Identities and News Consumption in Election Times" (PINCET; Bach et al., 2023). Members from an opt-in online panel in Germany provided by Respondi/Bilendi were invited through a survey-router system to participate in several waves of a web survey between August 30 and December 16, 2021. Panel members were eligible for the study if they were aged 18 years and older, lived in Germany, and were eligible to vote in the 2021 German federal election. Quotas for gender, age, and state were employed to generate a sample that aligned in these variables with the German general population.

To collect DBD from their devices, all panel members in this study had agreed to install a browser plug-in on their personal computers and/or download a research app on their mobile devices. The tracking technology, developed by Wakoopa, was provided by the online panel, and all participants had the technology already installed prior to the participation in this study. Each time a panel member visited a website, the browser plug-in captured the URL of the website, the domain, and the date, time, and duration of the visit. On mobile devices, the research app captured similar information about website visits from the device's native browser, that is, Chrome on Android devices and Safari on Apple devices. In addition, the research app captured the names of the apps that panel members used on their smartphone, including the date, time, and duration of any instance of use. Of all panel members who were invited to the DBD collection by the panel provider, approximately 30% allowed data collection on at least one device. In addition to incentives for survey participation, panel members who provided DBD received an incentive of €1 per month for data collection on a personal computer and €2 per month for data collection on a mobile device. If no longer interested in participating in the DBD collection, panel members could opt-out of the data collection or pause it temporarily at any time. In accordance with the EU General Data Protection Regulation (GDPR), panel members could also ask the sample provider to delete all their records. All DBD and survey data were provided to us in pseudonymized and de-identified form.

In this paper, we use data from $N = 1204$ smartphone users who completed the wave 1 survey of the PINCET project between August 30 and September 7, 2021 and had downloaded the research app on their smartphone. Participants have a median age of 45.5 years, 48.5% are female or diverse and 51.5% are male, and 23.9% have a college degree while 22.5% have a high school degree and 53.6% are without high school degree. For each participant, we use the app tracking data that were collected by the panel provider prior to their wave 1 survey interview between July 15 and September 7, 2021. During this period, participants provided DBD for a median of 47 days, with a range of 1–55 days. A total of 1153 participants (95.8%) installed the app on one smartphone and 51 participants (4.2%) on two smartphones, usually on one personal and one work-related device. The research app was available for Android and iOS operating systems, but the large majority of tracked devices are Android smartphones (88.9%; $n = 1116$) as opposed to iPhones (11.1%; $n = 139$).

Measures

For both the DBD and survey data, we create indicators of the amount, the variety, and the activities of smartphone use, reflecting the three dimensions of digital technology use that are

typically distinguished in digital divide research (Blank & Groselj, 2014). To measure *amount of use* in the survey, participants were asked to self-report the time spent using the smartphone on an ordinary day in an open answer box (in hours and minutes[1]). To measure *activities of use*, participants were asked to indicate the types of activities carried out on their smartphone in a check-all-that-apply question. The activities include (1) making and receiving phone calls, (2) using messenger services, (3) visiting websites, (4) sending and/or reading emails, (5) taking photos, (6) using social media, (7) shopping, (8) online banking, (9) using location-based apps, (10) playing games, (11) listening to music or watching videos, and (12) health and/or fitness tracking. In a separate question, they were also asked to self-report the frequency of (13) reading, listening, or watching the news on their smartphone. For each of the 13 activities, we create variables indicating whether the participants engage in the respective activities (coded as yes, no). Finally, to measure *variety of use*, the number of different activities carried out on the smartphone was summed up for each participant, ranging from 1 to 13. The original and translated questions are shown in Table S1 in the Online Appendix.

For the DBD, we create measures of smartphone use that mirror the survey-based measures. To measure *amount of use*, the time spent on all apps across all tracked smartphones during the data collection period were aggregated for each participant. The aggregated time was then divided by the number of days for which the participants' devices were tracked to create a measure on the day level. To measure *activities of use*, 13 variables were created indicating whether the participants engage in the smartphone activities that were also measured in the survey at least once during the data collection period (coded as yes, no). For the classification of apps into types of activities, the original categories provided by the app stores were used as the starting point and were refined through manual coding[2]. Finally, to measure *variety of use*, the number of different activities carried out on the smartphone was summed up for each participant, ranging from 0 to 13.

Descriptive statistics for the survey- and DBD-based measures are shown in Table 1. On the aggregate level, *amount of use* is considerably higher in the survey data (mean: 229.1 min; median: 180.0 min) than in the DBD (mean: 112.1 min; median: 84.8 min) whereas *variety of use* is almost completely aligned between the survey data (mean: 9.1 activities; median: 10.0 activities) and the DBD (mean: 8.4 activities; median: 9.0 activities). For *activities of use*, the level of alignment varies by type of activity. The majority of activities are closely aligned between the two data sources, including messenger services, emails, social media, online banking, GPS, and games. Other activities are reported by substantially more respondents in the survey, including phone calls, browsing websites, photos, or news, or underreported, including shopping, music or videos, and health or fitness.

As correlates of smartphone use, sociodemographic variables were collected in the survey, including age (in years), gender (male vs. female, diverse), and educational attainment (no high school degree, high school degree, college degree).

Analysis Strategy

To compare the alignment of DBD- and survey-based measures of smartphone use on the individual level, we create three types of variables following Araujo et al. (2017).

Absolute Error. For the continuous variables *amount of use* and *variety of use*, we subtract the DBD-based values from the survey-based values for each respondent and take the absolute value of the difference. For the dichotomous variables reflecting *activities of use*, we code whether the DBD-based and survey-based values are different (1 = different and 0 = not different).

Table 1. Descriptive Statistics.

	Survey data	Digital behavioral data
Amount of use	min	min
Mean	229.1	112.1
Median	180.0	84.8
SD	214.4	105.5
Variety of use	#	#
Mean	9.1	8.4
Median	10.0	9.0
SD	3.0	3.2
Activities of use	%	%
Phone calls	89.7	67.4
Messenger services	90.9	93.0
Browsing websites	78.5	14.7
Emails	83.3	81.3
Photos	84.6	71.8
Social media	66.4	71.4
Shopping	54.5	78.5
Online banking	65.7	66.4
GPS	67.2	63.9
Games	47.4	52.6
Music or videos	53.7	77.7
Health or fitness	35.0	60.5
News	94.5	36.1

Underreporting. Participants are considered to underreport a behavior if the behavior is observed in the DBD but not indicated in the survey. For the continuous variables *amount of use* and *variety of use*, we subtract the survey-based values from the DBD-based values for each participant. Instances in which participants report correctly or overreport their behavior in the survey are set as 0 for that variable. For the dichotomous variables reflecting *activities of use*, we code whether the behavior is observed in the DBD but not indicated in the survey (1 = underreporting and 0 = no underreporting).

Overreporting. Participants are considered to overreport a behavior if the behavior is indicated in the survey but not observed in the DBD. For the continuous variables *amount of use* and *variety of use*, we subtract the DBD-based values from the survey-based values for each participant. Instances in which participants report correctly or underreport their behavior in the survey are set as 0 for that variable. For the dichotomous variables reflecting *activities of use*, we code whether the behavior is indicated in the survey but not observed in the DBD (1 = overreporting and 0 = no overreporting).

To examine whether absolute error, underreporting, and overreporting are systematically related to sociodemographic characteristics, we estimate a series of regression models on age, gender, and educational attainment. We fit OLS regressions for the continuous variables *amount of use* and *variety of use*, and logistic regressions for the dichotomous variables reflecting *activities of use*. For the visualization of the regression results, we use the R ggplot2 package, version 3.4.1 (Wickham, 2016).

Finally, to create **typologies of smartphone use** based on the DBD and survey data, we conduct a latent class analysis (LCA) separately for each method of data collection. LCA is a clustering method for identifying unobserved classes in a population from a set of observed categorical indicators (McCutcheon, 1987). Participants are assigned to the different classes based on their similarity in the indicator variables. We use the 15 variables reflecting the three usage dimensions (*amount of use, variety of use*, and 13 variables reflecting *activities of use*) in the models. For the continuous variables *amount of use*, and *variety of use*, the values were coded into categories (below median vs. equal or above median), separately for the DBD and survey data. To estimate the latent class models, we use the R poLCA package, version 1.6.0.1 (Linzer & Lewis, 2011). For each model, we vary the number of classes from two to 10 and compute model fit criteria, including the log likelihood (LL), the Akaike information criterion (AIC), and the Bayesian information criterion (BIC), to select the best-fitting model, with lower values indicating a better model fit (Nylund et al., 2007). We also report the size and percentage of the smallest class. The data preparation and analysis were conducted in R, version 4.2.2 (R Core Team, 2022).

Results

We first examine to what extent the DBD- and survey-based measures of smartphone use align with each other on the individual level (Table 2). Generally, the individual-level alignment mirrors the patterns found on the aggregate level. For *amount of use*, the mean absolute error (AE) is 157.3 min (SD = 191.0) which is mostly due to overreporting. On average, participants overreport their daily smartphone use by 137.1 min (SD = 198.0) and only underreport their use by 20.1 min

Table 2. Comparison of Survey-Based and DBD-Based Measures of Smartphone Use.

	Absolute error	Underreporting	Overreporting
Amount of use	Min	Min	Min
Mean	157.3	20.1	137.1
Median	96.3	0.0	67.4
SD	191.0	52.8	198.0
Variety of use	#	#	#
Mean	3.1	1.1	1.9
Median	2.0	0.0	0.0
SD	2.6	1.9	2.7
Activities of use	%	%	%
Phone calls	34.7	6.2	28.5
Messenger services	13.5	7.8	5.7
Browsing websites	67.8	2.0	65.8
Emails	23.3	10.6	12.6
Photos	34.9	11.0	23.8
Social media	22.3	13.7	8.6
Shopping	40.0	32.0	8.0
Online banking	31.9	16.3	15.6
GPS	40.0	18.4	21.7
Games	28.6	16.9	11.7
Music or videos	39.4	31.7	7.6
Health or fitness	46.5	36.0	10.5
News	59.7	0.7	59.1

(SD = 52.8). For *variety of use*, the mean AE is 3.1 activities (SD = 2.6) which is due to both under- and overreporting. On average, participants underreport their variety of use by 1.1 activities (SD = 1.9) and overreport their use by 1.9 activities (SD = 2.7). For *activities of use*, the individual-level alignment varies considerably by type of activity. The largest differences between DBD- and survey-based measures are found for browsing websites (AE: 67.8%), news (AE: 59.7%), and health or fitness (AE: 46.5%). For browsing websites, the differences are mostly due to over-reporting (65.8%) rather than underreporting (2.0%). Similarly, the differences for news are largely driven by overreporting (59.1%) as opposed to underreporting (.7%). For health or fitness, in turn, underreporting is more prevalent (36.0%) than overreporting (10.5%). The closest alignment between DBD- and survey-based measures can be found for messenger services (AE: 13.5%), social media (AE: 22.3%), and emails (AE: 23.3%); any differences in these activities are driven by both under- and overreporting. Among the remaining activities with a medium level of alignment between DBD and survey data, overreporting is more prevalent for phone calls and photos whereas underreporting is more prevalent for shopping and music or videos. Finally, the differences for online banking, GPS, and games are evenly driven by under- and overreporting.

We next study to what extent the individual-level alignment is systematically related to participants' sociodemographic characteristics. Figures 1–5 show the results from regression models of absolute error, underreporting, and overreporting for *amount of use*, *variety of use*, and *activities of use*, with detailed model results shown in Tables S2–S6 in the Online Appendix. Predictors for which the horizontal lines do not cross the dashed zero-line are statistically significant at the 5%-level. We find that age, educational attainment, and gender are significantly

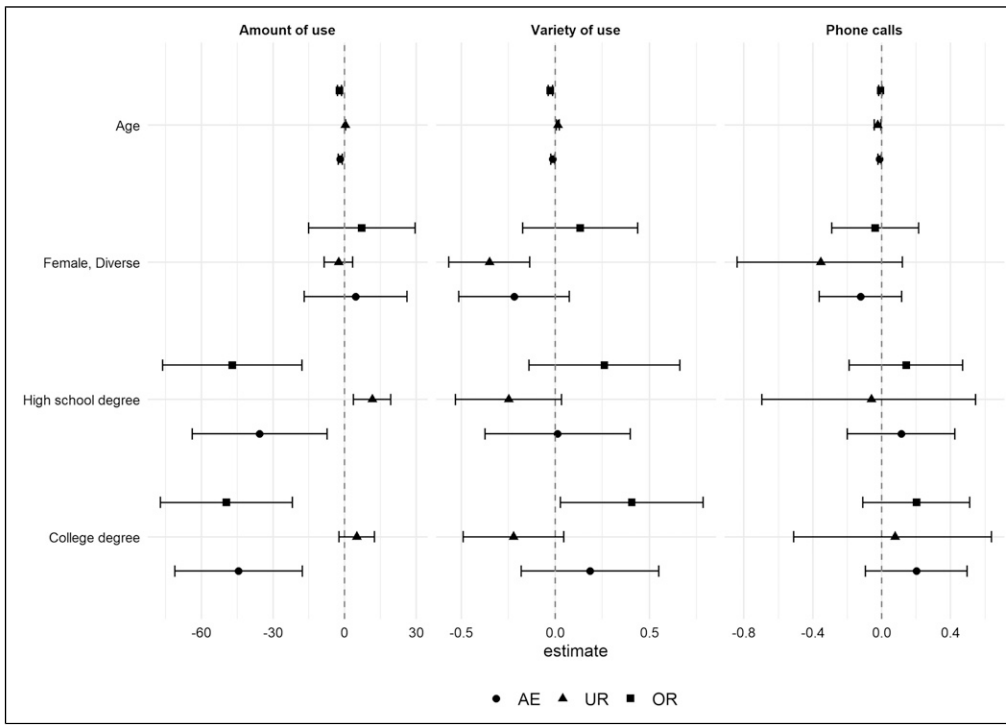

Figure 1. Coefficients (points) and 95% confidence intervals (lines) from regression models of absolute error (AE), underreporting (UR), and overreporting (OR) of smartphone use (amount of use, variety of use, and phone calls) on sociodemographic characteristics.

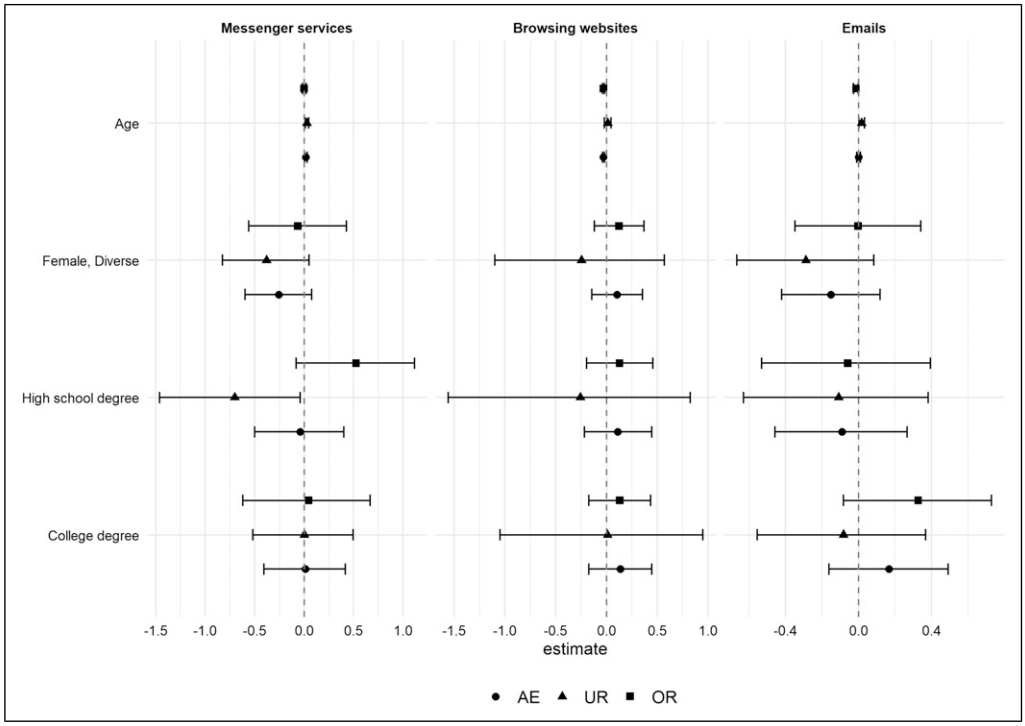

Figure 2. Coefficients (points) and 95% confidence intervals (lines) from regression models of absolute error (AE), underreporting (UR), and overreporting (OR) of smartphone use (messenger services, browsing websites, and emails) on sociodemographic characteristics.

correlated with the alignment between DBD and survey data for several dimensions of smartphone use. For *amount of use,* the absolute error and the level of overreporting significantly decrease with age. Similarly, for *variety of use*, the absolute error and the level of overreporting significantly decrease with age, although the level of underreporting significantly increases with age. For *activities of use*, the effect of age on the alignment between DBD and survey data varies by type of activity. Age has a significant negative effect on the absolute error and the level of overreporting for browsing websites, the level of overreporting for GPS, and the absolute error and the level of underreporting for phone calls. However, age has a significant positive effect on the absolute error and the level of underreporting for messenger services, shopping, online banking, and music or videos. The findings are mixed for emails and news where the level of overreporting significantly decreases with age, but the level of underreporting significantly increases with age.

We also find that educational attainment is significantly correlated with the alignment between DBD and survey data for all three dimensions of smartphone use. For *amount of use*, participants with a high school or college degree have a significantly smaller average error and lower level of overreporting than those without high school degree. However, participants with a high school degree have a significantly higher level of underreporting for *amount of use* than participants without high school degree. For *variety of use*, the alignment between the data sources decreases with the level of educational attainment. Participants with a college degree have a significantly higher level of overreporting than those without high school degree. For *activities of use*, educational attainment has a significant but mixed effect on few of the activities. Participants with a higher level of educational attainment are significantly more likely to overreport photos but

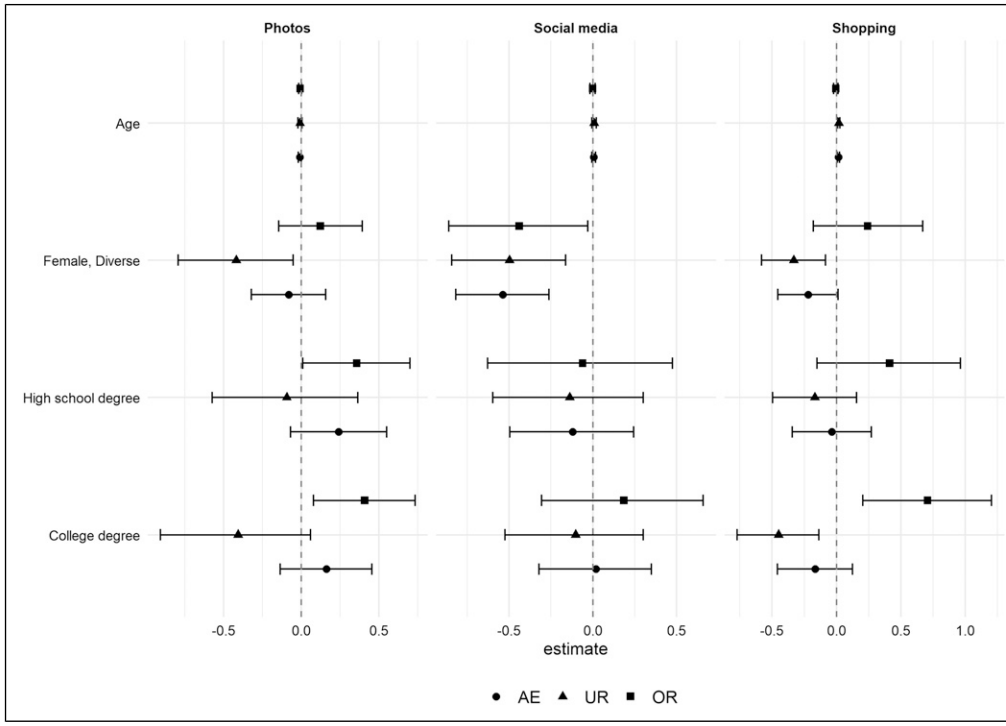

Figure 3. Coefficients (points) and 95% confidence intervals (lines) from regression models of absolute error (AE), underreporting (UR), and overreporting (OR) of smartphone use (photos, social media, and shopping) on sociodemographic characteristics.

significantly less likely to underreport GPS. In addition, participants with a college degree are significantly less likely to underreport shopping and also more likely to overreport shopping than those without high school degree.

Gender is significantly correlated with the alignment between the data sources for two of the usage dimensions. For *variety of use*, the level of underreporting is significantly lower for female or diverse participants compared to male participants. For *activities of use*, female or diverse participants have a significantly lower level of underreporting for photos, social media, shopping, and GPS, and a significantly lower level of overreporting for social media. However, female or diverse participants have a significantly higher absolute error and level of overreporting for news than male participants.

Finally, we investigate which types of smartphone users can be identified in the DBD and survey data by conducting LCAs. Varying the number of classes from two to 10 shows that the BIC reaches a minimum at the three-class model for the survey data, with a LL and AIC that do not decrease substantially as more classes are included in the model (Table 3). The three-class model also results in classes of a reasonable size, with the smallest class containing 156 participants (13% of the overall sample). We therefore select the three-class solution for the survey data. For the DBD, in turn, varying the number of classes from two to 10 shows that the BIC reaches a minimum at the four-class model. This model similarly results in classes of a reasonable size, with the smallest class containing 142 participants (12% of the overall sample), and we therefore select the four-class solution for the DBD.

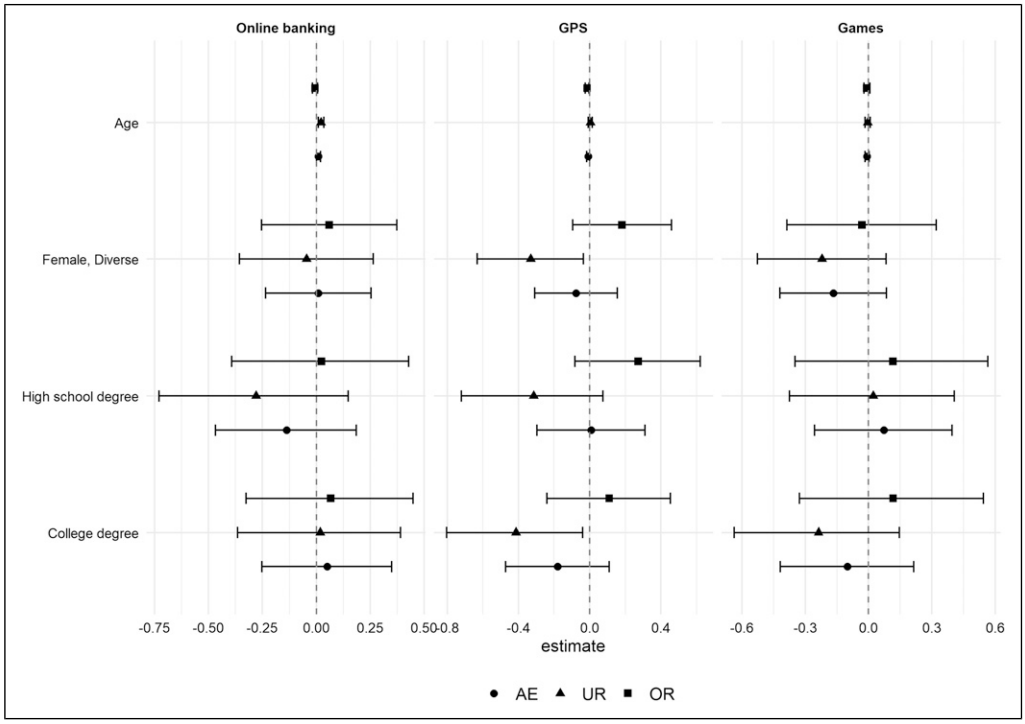

Figure 4. Coefficients (points) and 95% confidence intervals (lines) from regression models of absolute error (AE), underreporting (UR), and overreporting (OR) of smartphone use (online banking, GPS, and games) on sociodemographic characteristics.

We next examine the composition of the latent classes. Table 4 shows the predictor variables by smartphone usage class for the survey data and the DBD. For the survey data, we describe the three usage types as follows.

Heavy users report using their smartphone for a large amount of time (equal or above median: 67%) and a large variety of activities (equal or above median: 100%). The majority engage in each of the 13 activities on their smartphone (each activity used by at least 55%). Heavy users constitute the largest usage group, with more than half of the sample (53%) categorized into this group.

Intermediate users report using their smartphone for a medium amount of time (below median: 59%) and a smaller variety of activities (below median: 100%). The majority engage in most of the 13 activities, but they are considerably less likely than heavy users to use their smartphone for activities, such as health or fitness tracking (14%), shopping (24%), listening to music or watching videos (31%), and playing games (32%). Intermediate users constitute the second largest usage group, with one third of the sample (34%) categorized into this group.

Light users report using their smartphone for a shorter amount of time (below median: 86%) and a smaller variety of activities (below median: 100%). They mainly use their device for news consumption (82%), making and receiving phone calls (73%), using messenger services (61%), and taking photos (41%). Light users constitute the smallest usage group, with 13% of the sample categorized into this group.

For the DBD, in turn, we describe the four usage types as follows.

Heavy users use their smartphone for a large amount of time (equal or above median: 68%) and a large variety of activities (equal or above median: 100%). The majority engage in most of the 13

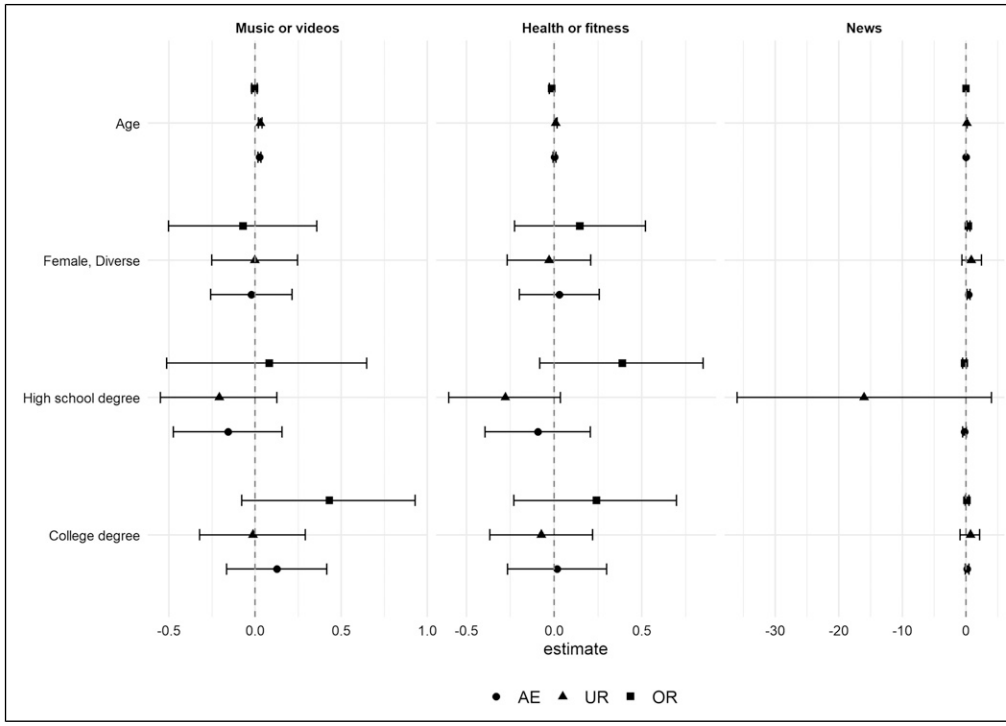

Figure 5. Coefficients (points) and 95% confidence intervals (lines) from regression models of absolute error (AE), underreporting (UR), and overreporting (OR) of smartphone use (music or videos, health or fitness, and news) on sociodemographic characteristics.

activities on their smartphone (each activity used by at least 52%, except for browsing websites with only 22%). Heavy users constitute the largest usage group, with more than half of the sample (56%) categorized into this group.

Intermediate social media users use their smartphone for a medium amount of time (below median: 57%) and a smaller variety of activities (below median: 100%). The most popular activities are using messenger services (97%), social media (92%), and shopping (83%). These users also commonly use their device for sending or reading emails (73%), listening to music or watching videos (72%), and online banking (53%). Intermediate social media users constitute 16% of the sample.

Intermediate phone call users use their smartphone for a medium amount of time (below median: 72%) and a smaller variety of activities (below median: 100%). The most popular activities are using messenger services (93%), making and receiving phone calls (84%), and taking photos (84%). These users also commonly use their device for sending or reading emails (70%), listening to music or watching videos (62%), and shopping (62%). Intermediate phone call users constitute 16% of the sample.

Light users use their smartphone for a shorter amount of time (below median: 98%) and a smaller variety of activities (below median: 100%). They mainly use their device for using messenger services (54%) and sending or reading emails (30%). Light users constitute 12% of the sample.

Overall, the LCAs based on survey data and DBD result in similar smartphone usage classes, but the DBD reveal more nuanced patterns of smartphone use. Heavy users and light users have a similar composition and size across both data sources, although light users differ in the types of

Table 3. Model Fit and Diagnostic Criteria for Two to Ten Classes of Smartphone Use.

	Survey data					Digital behavioral data				
	Model fit criteria			Diagnostic criteria		Model fit criteria			Diagnostic criteria	
# Classes	LL	AIC	BIC	Smallest class (*n*)	Smallest class (%)	LL	AIC	BIC	Smallest class (*n*)	Smallest class (%)
2	−8,056.01	16,174.02	16,331.92	568	47	−8,410.12	16,882.24	17,040.13	524	44
3	**−7,925.39**	**15,944.78**	**16,184.17**	**156**	**13**	−8,229.07	16,552.13	16,791.52	122	10
4	−7,906.86	15,939.72	16,260.61	38	3	**−8,128.19**	**16,382.37**	**16,703.25**	**142**	**12**
5	−7,876.83	15,911.65	16,314.03	40	3	−8,077.90	16,313.80	16,716.18	142	12
6	−7,848.83	15,887.65	16,371.52	105	9	−8,055.55	16,301.09	16,784.96	77	6
7	−7,831.72	15,885.45	16,450.82	79	7	−8,031.05	16,284.10	16,849.47	71	6
8	−7,818.91	15,891.83	16,538.69	18	1	−8,013.87	16,281.74	16,928.60	41	3
9	−7,789.18	15,864.35	16,592.71	39	3	−7,986.89	16,259.77	16,988.13	32	3
10	−7,769.85	15,857.70	16,667.55	13	1	−7,981.69	16,281.38	17,091.23	17	1

Note. LL = log likelihood; AIC = Akaike information criterion; BIC = Bayesian information criterion. The rows in bold indicate the latent class models chosen for the analysis.

Table 4. Predictor Variables by Class of Smartphone Use.

	Survey data			Digital behavioral data			
Variables	Heavy users %	Intermediate users %	Light users %	Heavy users %	Intermediate social media users %	Intermediate phone call users %	Light users %
Class size	53	34	13	56	16	16	12
Amount of use							
Below median	33	59	86	32	57	72	98
Equal or above median	67	41	14	68	43	28	2
Variety of use							
Below median	0	100	100	0	100	100	100
Equal or above median	100	0	0	100	0	0	0
Activities of use							
Phone calls	97	85	73	92	5	84	9
Messenger services	99	90	61	100	97	93	54
Browsing websites	98	71	18	22	5	4	5
Emails	99	78	31	98	73	70	30
Photos	97	81	41	96	16	84	14
Social media	89	51	13	84	92	44	23
Shopping	88	24	0	96	83	62	13
Online banking	88	53	11	89	53	39	13
GPS	93	48	13	92	19	49	9
Games	65	32	16	68	36	36	21
Music or videos	80	31	6	95	72	62	24
Health or fitness	55	14	10	86	37	31	8
News	98	93	82	52	14	22	9

activities that are most frequently used on their smartphones. Intermediate users identified in the survey data, however, are reflected by two classes of smartphone users in the DBD that differ in their activities of use. While intermediate phone call users engage in more classical types of mobile phone activities, such as making and receiving phone calls and taking photos, intermediate social media users are more likely to engage in Internet-based activities, such as social media, shopping, and online banking.

Finally, we investigate to what extent the classification into smartphone usage groups aligns on the individual level (Table 5). Whereas the aggregate-level comparison shows that the groups have

Table 5. Alignment Between Survey-Based and DBD-Based Classes of Smartphone Use.

%	Survey data		
Digital behavioral data	Heavy users	Intermediate users	Light users
Heavy users	32	19	5
Intermediate social media users	11	4	1
Intermediate phone call users	5	7	4
Light users	5	5	2

a similar size, the results from the individual-level analysis suggest that there is a rather small overlap between the usage groups identified in the two data sources. Less than half (45%) of the participants are classified into the same group of smartphone users in both survey data and DBD. The alignment is much closer for heavy users than for the other usage groups: 32% of participants are classified as heavy users in both data sources, although an additional 11% are classified as heavy users in the survey data and as intermediate social media users in the DBD. In contrast, only 11% of participants are classified as intermediate users in the survey data and as either intermediate social media users or intermediate phone call users in the DBD. Light users have the smallest alignment: Only 2% of participants are classified as light users in both data sources while an additional 10% are classified as light users in the survey data but in a different usage group in the DBD.

Discussion

While digital technology use and skills have mostly been measured with surveys, DBD that are passively collected from individuals' digital devices might serve as an additional method of measuring technology usage patterns in a more unobtrusive and detailed way. In this paper, we contribute to the growing body of research about differences in DBD and survey data for measuring smartphone use. Based on a sample of smartphone users in Germany who completed a survey and had a tracking app installed on their smartphone, we examine to what extent the DBD- and survey-based usage measures align with each other. In addition, we investigate whether the DBD and survey data lead to different typologies of smartphone use.

The results show that the level of alignment between DBD and survey data for measuring smartphone use varies by usage dimension. Whereas *amount of use* is considerably overreported in the survey data compared to the DBD, *variety of use* aligns more closely across the two data sources. For *activities of use*, the level of alignment differs by type of activity. On the one hand, there are activities that align relatively closely across DBD and survey data, such as using messenger services, social media, and sending or reading emails. On the other hand, there are activities that are either considerably overreported in the survey, such as browsing websites, news consumption, and making or receiving phone calls, or underreported, such as shopping, listening to music or watching videos, and health or fitness tracking. Overall, these findings are in line with previous research showing that the correlations between self-reported and tracked measures of digital technology use are generally small to moderate (e.g., Parry et al., 2021). In addition, we find that the level of alignment is systematically related to participants' sociodemographic characteristics. Differences in *amount of use* decrease with age whereas differences in *variety of use* decrease with age and educational attainment and are smaller for female or diverse participants compared to male participants. Differences in *activities of use* are also smaller for female or diverse participants for several types of activities, except for news consumption where differences are smaller for male participants. The effects of age and educational attainment on the alignment

between DBD and survey data in *activities of use* are rather mixed. Finally, the DBD and survey data lead to similar typologies of smartphone use, although the DBD-based typology is more nuanced, with a larger number of classes. Whereas the classes are similar in size and composition on the aggregate level, the overlap between the two typologies on the individual level is rather small.

There are several potential explanations for the differences between DBD and survey data in measuring smartphone use. On the one hand, the differences could have arisen due to measurement error in the survey-based self-reports (Tourangeau et al., 2000). Participants might not have been able to recall all activities that they usually carry out on their smartphone, or the phrasing of the survey questions might not have prompted them to think of all instances of use, leading to an underreporting of certain types of smartphone activities. For example, the question about news consumption does not further specify the type of news: While recalling the use of news apps on their smartphone, respondents might not have considered news consumption within other apps, such as mobile browser or social media apps. An additional limitation of the questions about activities of smartphone use is that they lack a specific time frame, which might have helped respondents to recall the types of activities carried out on their smartphone. In addition, social desirability might have been at play, leading to a misreporting of certain activities of use, such as an overreporting of news consumption. Participants might have also had difficulties recalling their amount of smartphone use on an ordinary day and reporting a continuous number in hours and minutes. We welcome future research that replicates our analysis with ordinal measures of amount of use, such as those employed in Blank and Groselj (2014).

On the other hand, errors in the DBD could have contributed to the observed differences (Bosch & Revilla, 2022; Keusch et al., 2023; Scharkow, 2016). Since only the name of the app is recorded and not the type of activity carried out within the app, the classification of apps into activities of use might have been erroneous, in particular for apps that can potentially serve numerous activities. For example, social media apps might be used for various activities, such as news consumption, messaging, or playing games. Future research might consider data collection approaches that allow monitoring within-app activities, such as data donation of social media data (Ohme et al., 2023). In addition, participants might not have installed the research app on all their smartphones or might have turned off the tracking due to privacy concerns, resulting in an underestimation of smartphone use. The participants' awareness of being observed might also have changed their smartphone usage behavior in the course of the study. Furthermore, it cannot be ruled out that the tracked smartphones were used by multiple people, in which case the participants' smartphone use is overestimated. Finally, although the research app passively collected DBD through the smartphone operating system log files, these data are not perfectly accurate and might also contain technical errors.

Another explanation why we see relatively little alignment between the self-reports and the DBD could be that the two data sources conceptually measure different dimensions of smartphone behavior. While self-reports measure perception of how frequently someone uses their smartphone and for what purpose, the DBD allow for a direct technical measure of actual behavior. Both aspects of the measurement could be relevant in a study, and researchers have to decide on a case-by-case basis which measure best operationalizes the concept of interest. For activities of use, for example, survey data could be better at measuring smartphone activities that can be carried out across multiple apps, such as news consumption or watching videos, whereas DBD could be better for activities carried out within distinct apps, such as making or receiving phone calls or taking photos. Given that both self-reports and DBD come with measurement error, approaches that incorporate both measures without assuming one of them being the gold standard might be the best fit for studies going forward, providing a more holistic view of smartphone use.

Declaration of Conflicting Interests

The author(s) declared no potential conflicts of interest with respect to the research, authorship, and/or publication of this article.

Funding

The authors disclosed receipt of the following financial support for the research, authorship, and/or publication of this article: Ruben Bach gratefully acknowledges financial support by the Eliteprogramme for Postdocs of the Baden-Württemberg Stiftung.

ORCID iDs

Alexander Wenz https://orcid.org/0000-0002-4621-2418
Florian Keusch https://orcid.org/0000-0003-1002-4092
Ruben L. Bach https://orcid.org/0000-0001-5690-2829

Supplemental Material

Supplemental material for this article is available online.

Data Availability Statement

The survey data are available from the GESIS Data Archive for the Social Sciences at https://doi.org/10.7802/2524. The digital behavioral data are available upon request from Ruben Bach at r.bach@uni-mannheim.de.

Notes

1. To account for outliers in the amount of smartphone use, values smaller than the first percentile (10 min) or larger than the 99th percentile (1200 min) were winsorized to the first percentile or 99th percentile values.
2. The scheme for classifying apps into categories and the analysis code are available at https://doi.org/10.17605/OSF.IO/RV4YX.

References

Andrews, S., Ellis, D. A., Shaw, H., & Piwek, L. (2015). Beyond self-report: Tools to compare estimated and real-world smartphone use. *PLoS One*, *10*(10), Article e0139004. https://doi.org/10.1371/journal.pone.0139004

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 and Measures*, *11*(3), 173–190. https://doi.org/10.1080/19312458.2017.1317337

Bach, R., Keusch, F., Areal, J., Pankowska, P., & Cernat, A. (2023). *Political identities and news consumption in election times (PINCET) (Data File Version 1.0.0)*. Cologne: GESIS Data Archive. https://doi.org/10.7802/2524

Bach, R. L., & Wenz, A. (2020). Studying health-related internet and mobile device use using web logs and smartphone records. *PLoS One*, *15*(6), Article e0234663. https://doi.org/10.1371/journal.pone.0234663

Barthel, M., Mitchell, A., Asare-Marfo, D., Kennedy, C., & Worden, K. (2020). *Measuring news consumption in a digital era*, Pew Research Center. https://www.pewresearch.org/journalism/2020/12/08/measuring-news-consumption-in-a-digital-era/

Baumgartner, S. E., Sumter, S. R., Petkevič, V., & 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

Blank, G., & Groselj, D. (2014). Dimensions of Internet use: Amount, variety, and types. *Information, Communication and Society*, *17*(4), 417–435. https://doi.org/10.1080/1369118X.2014.889189

Boase, J., & Ling, R. (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

Bosch, O. J., & Revilla, M. (2022). When survey science met web tracking: Presenting an error framework for metered data. *Journal of the Royal Statistical Society Series A: Statistics in Society*, *185*(Suppl 2), S408–S436. https://doi.org/10.1111/rssa.12956

Büchi, M., Just, N., & Latzer, M. (2016). Modeling the second-level digital divide: A five-country study of social differences in internet use. *New Media and Society*, *18*(11), 2703–2722. https://doi.org/10.1177/1461444815604154

Chen, H., Zhang, L., Chu, X., & Yan, B. (2019). Smartphone customer segmentation based on the usage pattern. *Advanced Engineering Informatics*, *42*, 101000. https://doi.org/10.1016/j.aei.2019.101000

Deng, T., Kanthawala, S., Meng, J., Peng, W., Kononova, A., Hao, Q., Zhang, Q., & David, P. (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*(August), 60–66. https://doi.org/10.1016/j.chb.2019.03.006

Ellis, D. A., Davidson, B. I., Shaw, H., & Geyer, K. (2019). Do smartphone usage scales predict behavior? *International Journal of Human-Computer Studies*, *130*, 86–92. https://doi.org/10.1016/j.ijhcs.2019.05.004

Festic, N., Büchi, M., & Latzer, M. (2021). How long and what for? Tracking a nationally representative sample to quantify internet use. *Journal of Quantitative Description: Digital Media*, *1*(April), 1–23. https://doi.org/10.51685/jqd.2021.018

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

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

Hamka, F., Bouwman, H., De Reuver, M., & Kroesen, M. (2014). Mobile customer segmentation based on smartphone measurement. *Telematics and Informatics*, *31*(2), 220–227. https://doi.org/10.1016/j.tele.2013.08.006

Harari, G. M., Lane, N. D., Wang, R., Crosier, B. S., Campbell, A. T., & Gosling, S. D. (2016). Using smartphones to collect behavioral data in psychological science: Opportunities, practical considerations, and challenges. *Perspectives on Psychological Science*, *11*(6), 838–854. https://doi.org/10.1177/1745691616650285

Harari, G. M., Müller, S. R., Stachl, C., Wang, R., Wang, W., Bühner, M., Rentfrow, P. J., Campbell, A. T., & Gosling, S. D. (2020). Sensing sociability: Individual differences in young adults' conversation, calling, texting, and app use behaviors in daily life. *Journal of Personality and Social Psychology*, *119*(1), 204–228. https://doi.org/10.1037/pspp0000245

Hargittai, E. (2005). Survey measures of web-oriented digital literacy. *Social Science Computer Review*, *23*(3), 371–379. https://doi.org/10.1177/0894439305275911

Jones-Jang, S. M., Heo, Y.-J., McKeever, R., Kim, J.-H., Moscowitz, L., & Moscowitz, D. (2020). Good news! Communication findings may be underestimated: Comparing effect sizes with self-reported and logged smartphone use data. *Journal of Computer-Mediated Communication*, *25*(5), 346–363. https://doi.org/10.1093/jcmc/zmaa009

Keusch, F., Bach, R., & Cernat, A. (2023). Reactivity in measuring sensitive online behavior. *Internet Research*, *33*(3), 1031–1052. https://doi.org/10.1108/INTR-01-2021-0053

Keusch, F., & Kreuter, F. (2022). Digital trace data: Modes of data collection, applications, and errors at a glance. In U. Engel, A. Quan-Haase, S. X. Liu, & L. Lyberg (Eds.), *Handbook of computational social science, volume 1: Theory, case studies and ethics* (pp. 100–118). Routledge.

Lee, Y., Park, I., Cho, S., & Choi, J. (2018). Smartphone user segmentation based on app usage sequence with neural networks. *Telematics and Informatics*, *35*(2), 329–339. https://doi.org/10.1016/j.tele.2017.12.007

Linzer, D. A., & Lewis, J. B. (2011). poLCA: An R package for polytomous variable latent class analysis. *Journal of Statistical Software*, *42*(10), 1–29. https://doi.org/10.18637/jss.v042.i10

Mahalingham, T., McEvoy, P. M., & Clarke, P. J. F. (2023). Assessing the validity of self-report social media use: Evidence of no relationship with objective smartphone use. *Computers in Human Behavior*, *140*(March), 107567. https://doi.org/10.1016/j.chb.2022.107567

McCutcheon, A. L. (1987). *Latent class analysis*. Sage Publications.

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

Nylund, K. L., Asparouhov, T., & Muthén, B. O. (2007). Deciding on the number of classes in latent class analysis and growth mixture modeling: A Monte Carlo simulation study. *Structural Equation Modeling*, *14*(4), 535–569. https://doi.org/10.1080/10705510701575396

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*, 1–18. https://doi.org/10.1080/19312458.2023.2181319

Ohme, J., Araujo, T., de Vreese, C., & 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. A., Davidson, B. I., Sewall, C. J. R., 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 Behaviour*, *5*(11), 1535–1547. https://doi.org/10.1038/s41562-021-01117-5

R Core Team. (2022). *R: A language and environment for statistical computing*, R Foundation for Statistical Computing. https://www.r-project.org/

Reisdorf, B. C., & Groselj, D. (2017). Internet (non-)use types and motivational access: Implications for digital inequalities research. *New Media and Society*, *19*(8), 1157–1176. https://doi.org/10.1177/1461444815621539

Revilla, M., Ochoa, C., & Loewe, G. (2017). Using passive data from a meter to complement survey data in order to study online behavior. *Social Science Computer Review*, *35*(4), 521–536. https://doi.org/10.1177/0894439316638457

Rice, R. E., Pearce, K. E., & Calderwood, K. J. (2023). W(h)ither the device divide? Changing relationships between personal computer or mobile device with online activities. *Mobile Media & Communication*, *11*(3), 484–506. https://doi.org/10.1177/20501579221142134

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

Schoedel, R., Pargent, F., Au, Q., Völkel, S. T., Schuwerk, T., Bühner, M., & Stachl, C. (2020). To challenge the morning lark and the night owl: Using smartphone sensing data to investigate day-night behaviour patterns. *European Journal of Personality*, *34*(5), 733–752. https://doi.org/10.1002/per.2258

Stier, S., Mangold, F., Scharkow, M., & Breuer, J. (2022). Post post-broadcast democracy? News exposure in the age of online intermediaries. *American Political Science Review*, *116*(2), 768–774. https://doi.org/10.1017/S0003055421001222

Sturgis, P., & Luff, R. (2021). The demise of the survey? A research note on trends in the use of survey data in the social sciences, 1939 to 2015. *International Journal of Social Research Methodology*, *24*(6), 691–696. https://doi.org/10.1080/13645579.2020.1844896

Tourangeau, R., Rips, L. J., & Rasinski, K. A. (2000). *The Psychology of Survey Response*. Cambridge University Press.

van Deursen, A. J. A. M., & van Dijk, J. A. G. M. (2014). The digital divide shifts to differences in usage. *New Media and Society*, *16*(3), 507–526. https://doi.org/10.1177/1461444813487959

van Deursen, A. J. A. M., Helsper, E. J., & Eynon, R. (2016). Development and validation of the internet skills scale (ISS). *Information, Communication and Society*, *19*(6), 804–823. https://doi.org/10.1080/1369118X.2015.1078834

Wenz, A., & Keusch, F. (2023). The second-level smartphone divide: A typology of smartphone use based on frequency of use, skills, and types of activities. *Mobile Media & Communication*, *11*(3), 459–483. https://doi.org/10.1177/20501579221140761

Wickham, H. (2016). *ggplot2: Elegant graphics for data analysis*. Springer-Verlag. https://ggplot2.tidyverse.org/

Zillien, N., & Hargittai, E. (2009). Digital distinction: Status-specific types of internet usage. *Social Science Quarterly*, *90*(2), 274–291. https://doi.org/10.1111/j.1540-6237.2009.00617.x

Author Biographies

Alexander Wenz is a Research Fellow at the Mannheim Centre for European Social Research at the University of Mannheim. His research investigates the quality of novel methods of data collection, with a focus on mobile web surveys, smartphone apps, wearable sensors, and digital behavioral data.

Florian Keusch is Professor of Social Data Science and Methodology at the School of Social Sciences at the University of Mannheim. He studies the quality of digital behavioral data from wearables, apps, and sensors and how they can complement survey data to better measure attitudes, behaviors, and social interactions.

Ruben L. Bach is a Research Fellow at the Mannheim Centre for European Social Research at the University of Mannheim. His research investigates novel data products, their applications in the social sciences, and methods of computational social science.