LOST IN TRACE-LATION

An Empirical Inquiry into the Role of Theoretical Interpretation, Granularity, and Scope of Digital Trace Data

Kyle Van Gaeveren

Doctoral dissertation submitted to obtain the degree of Doctor of Communication Sciences

Faculty of Political and Social Sciences
Department of Communication Sciences
imec-mict-UGent
Ghent University

Academic year 2025–2026

To refer or cite this work: Van Gaeveren, K. (2026). Lost in Trace-lation: An Empirical Inquiry into the Role of Theoretical Interpretation, Granularity, and Scope of Digital Trace Data. Ghent University. Faculty of Political and Social Sciences, Ghent, Belgium.

Examination Board

Chairman	
Prof. Dr. Stijn Joye	Department of Communication Sciences, Ghent University
Supervisors	
Prof. Dr. Mariek Vanden Abeele	Department of Communication Sciences, Ghent University
Prof. Dr. Lieven De Marez	Department of Communication Sciences, Ghent University
Reading Committee	
Prof. Dr. Els Clays	Department of Public Health and Primary Care, Ghent University
Dr. Yannick Vander Zwalmen	Department of Experimental Clinical and Health Psychology, Ghent University
Dr. Susanne Baumgartner	Amsterdam School of Communication Research, University of Amsterdam
Dr. Doug Parry	Department of Communication Science, Vrije Universiteit Amsterdam
Dr. Aurelio Fernández	Instituto Cultura y Sociedad, Universidad de Navarra

Acknowledgements

I'll start this dissertation with some words of thanks for those who helped me cross the finish line.

Lieven, thank you. Thank you for bringing me on board at mict, allowing me to take my time in getting to know digital trace data, letting me practice my skills and always responding with enthusiasm when I'd present you with some new plot, insight or dashboard. Thank you for greenlighting the shift to Mariek's ERC and this PhD trajectory.

Mariek, thank you. Thank you for considering me as your *data guy* for the team and giving me the confidence to trust that all other aspects of this PhD would also work out in the end. Thank you for your infectious positivity and scientific curiosity. Thank you for your patience and understanding at times when I (felt I) was lagging behind. Thank you for making time when you'd clearly forgotten my submission deadline, frantically scanning your overflowing calendar, but a few days later telling me this work was '*defendable*'.

Stephen, thank you. Thank you for being a rational, critical and honest sounding board anytime I'd feel blocked or overwhelmed by an avalanche of potential features and models, or by dreaded blank pages.

David & Sara, thank you. Thank you for giving me a sense of belonging in a team. Thank you for being there, sharing so much of this journey together, and providing me with two of the best examples of how to do all of this.

To all my mict-colleagues and ex-colleagues, thank you. Thank you for being such a diverse group, with some of the nicest, smartest, funniest people in the entire… 4th floor of De Krook. Thank you for the shared meals – with or without Tupperware –, drinks, teambuilding, quizzes, and warm welcomes whenever I passed by with Remi.

To anyone whom I've ever kicked a ball around with, thank you. Thank you for providing me with some much needed disconnection and distraction. Thank you in particular to **Stijn & Dries** for putting their time and effort into keeping AC/RAK going, and to all Ploegmoats who I've stubbornly convinced to keep their Friday evenings free. Thanks.

To my friends, thank you. Thanks for asking how this PhD was going, and thank you for *not* asking how this PhD was going. Thanks for reaching out and checking in.

To my parents, all three sets of them.

Mama & Alain, bedankt om zo'n goede zorg te dragen voor Remi, Hannah en mezelf. De tijd die ik niet altijd bij hen kon zijn, hebben jullie in een veelvoud opgevangen. Bedankt voor al jullie steun, zeker in tijden waar mijn schouders niet breed genoeg waren, mijn hoofd te veel bij dit doctoraat was, en mijn twee handen te linkshandig waren. Zonder jullie hulp hadden we geen huis om "thuis" te noemen en was dit nooit gelukt.

Papa & Klara, bedankt voor al jullie liefde en steun. Bedankt om er telkens te zijn als ik terug thuis kwam in Oostende. Bedankt om jullie nieuwsgierigheid en kritische ingesteldheid te delen. Ook al is de afstand nu wat groter, ik ben heel dankbaar jullie (digitaal) dichtbij te hebben.

Katlijne en Leon, bedankt om steeds de hele familie voorop te plaatsen. Bedankt voor alle gedeelde maaltijden en momenten, en jullie zorg voor Remi, ook op momenten waarop jullie heel wat andere zorgen hebben.

To **Hannah**, thank you. Thank you for repeatedly pointing out that a PhD isn't always the best idea for everyone, and that your experiences as a doctor had made that quite clear to you. Sorry, but, well I guess I was more stubborn. Nevertheless, thank you for giving me the belief that I could handle this. Thank you for threatening to call “Diene Lieven” or Mariek whenever I felt overworked. Thank you for being there as my wife, as Remi’s mom, as my love. Ik zie je graag.

To **Remi**, thank you. For your love, for making me laugh and smile every single day. Thank you for putting this all in perspective.

Table of Contents

List of Tables

List of Figures

General Introduction

There is ample debate in society over the impact of (social media) screen time on well-being (Ferguson et al., 2022; Twenge & Farley, 2021; Valkenburg, 2022). What fuels this debate is that studies on the topic provide conflicting results. For instance, Twenge and colleagues identified screen time as a direct predictor of mental health problems (e.g., Twenge et al., 2018), while others find positive dynamics between smartphone use and well-being (Marciano et al., 2022; Marciano & Viswanath, 2023). In a recent umbrella review synthesizing systematic reviews and meta-analyses on this association, Valkenburg and colleagues (2022) concluded that the extant evidence is indeed mixed, with methodological shortcomings stated as a key reason.

In particular, the field has long relied on self-reported measures of screen time, whose validity has been repeatedly questioned: such reports are inaccurate (Hodes & Thomas, 2021; Molaib et al., 2025; Ohme et al., 2021; Parry et al., 2021; Tkaczyk et al., 2024; Verbeij et al., 2021) and systematically biased (Sewall et al., 2020). For instance, Sewall and colleagues (2020) found that the accuracy of self-reports is associated with well-being and usage amount.

The concerns over the validity of self-reported screen time have motivated a noticeable paradigm shift over the past decade (Ellis et al., 2019; Parry et al., 2021), in which researchers are increasingly moving away from retrospective self-reports and towards the use of behavioral data (e.g., Buda et al., 2023; de Segovia Vicente et al., 2024; Elmer et al., 2025; große Deters & Schoedel, 2024; Rodman et al., 2024; Schenkel et al., 2024).

The collected behavioral data, also commonly referred to as digital trace data (Aalbers, 2023), typically fall within the realm of 'big data', leveraging thousands of data points per participant. Although these data often require extensive cleaning and pre-processing, their use enables more sophisticated methods for feature extraction, analysis and modeling (Bjerre-Nielsen & Glavind, 2022; Boeschoten et al., 2022; Parry & Toth, 2025; Peters et al., 2024; Saito et al., 2015). As such, the paradigm shift towards the use of passively sensed device data can be argued to also be part of the 'computational turn' in the broader field of communication, and even the social sciences (Berry, 2011; Es et al., 2021; Hepp et al., 2021).

Within this computational turn, the smartphone has become a main device of interest when it comes to the study of screen time and well-being. Smartphone trace data offer higher temporal resolution and reduce reliance on participants' retrospective recall, enabling data collection within the context of daily life (Harari et al., 2017; Harari & Gosling, 2023). When combined with (mobile) Experience Sampling Methods (ESM), this provides the opportunity to carry out more precise analyses, taking within-person differences, temporal dimensions, and even momentary associations into account (Elmer et al., 2025; Harari & Gosling, 2023; Keusch & Conrad, 2022; Klingelhoefer et al., 2025).

The methods used to collect behavioral data are varied. For instance, researchers have used platform-specific (e.g., Facebook, Twitter/X) API (Application programming interface) approaches, Data Download Packages (DDPs) from those same platforms, screenshots from smartphone battery usage or from Screen Time (iPhone) or Digital Wellbeing (Android) tools, network traffic monitoring, and device-level logging. While these methods all provide objective measures, they differ in granularity and the type of research question they can help answer best.

To address the research questions posed in this dissertation, I specifically focus on utilizing **event logs** collected from both smartphones (Android) and personal computers (Windows/ Mac). These event logs include timestamped data on app launches, screen activation, and incoming notifications. The decision to use event logs stems from the context of the broader project in which the empirical chapters of this dissertation are situated. This project, called project DISCONNECT, aimed to 'unravel the science behind digital well-being'. The project, which is still ongoing, places the "always-on" nature of our digital technologies at the center. The fine-grained timing of these event logs was expected to support the modelling of short-term and within-person processes. Yet, the project was also conceived as an opportunity to critically examine whether such data truly deliver on their promise, or whether the shift from self-report to behavioral measurement introduces new challenges of its own.

This critical examination is warranted, as despite their advantages, behavioral data come with their own limitations, or, more precisely, with a set of validity considerations that differ from, but do not necessarily improve upon, those of self-reports. This dissertation focuses on three such challenges. The first concerns that of operationalization, or – in reverse – the theoretical interpretation of digital trace data: although digital trace data provide precise records of digital behavior, they lack information on aspects such as who performed the behavior, in which state or context, or what content was consumed. This means they do not speak for themselves, but sometimes require theoretical interpretation to become meaningful. For instance, Birk and Samuel (2020) argue in their sociological analysis of the use of trace data for digital phenotyping, that the lack of information on these aspects may lead to misinterpretations: An individual who spends hours voice calling, for example, may be interpreted as 'being highly social' while in reality the participant may have been on the line with a helpdesk. In short, without insights into how individuals experience or interpret digital behaviors, digital traces risk being used as proxies of cognitive-behavioral states that they fail to reliably represent. This raises the question of how such indicators can be meaningfully linked to subjective experiences of what they are supposed to represent.

The second challenge concerns granularity: At which level of detail should behavioral indicators be operationalized. Many studies have relied on daily or weekly aggregated metrics at the device level (e.g., total daily smartphone duration), which risks obscuring short, repeated or fragmented interactions which may increasingly characterize everyday digital behavior. Although emerging work suggests that more granular and dynamic indicators may offer greater inferential value (Rodman et al., 2024; Siebers et al., 2024), it remains unclear whether this increased complexity translates into meaningfully stronger or more theoretically coherent associations with well-being outcomes. This dissertation therefore treats granularity as an empirical question to be tested.

The third challenge concerns measurement scope: current research often employs a narrow measurement scope by focusing almost exclusively on smartphones as proxies for digital media use more broadly, rarely considering trace data from alternative digital devices such as laptops, tablets or televisions – leaving the question about 'total screen time' underspecified. Although this issue is often mentioned in discussions of study limitations (e.g., Johannes, Nguyen, et al., 2021; Mahalingham et al., 2023; Verbeij et al., 2021), its implications for the validity of empirical claims remain underexplored. Excluding other digital devices should therefore not be seen as a minor gap, but rather as a potential source of systematic bias in estimating both the extent of digital engagement and its association with well-being.

These three challenges show that advances in behavioral measurement may not automatically translate into new empirical or theoretical insight. Instead, they show the need to align theoretical interpretation, granularity and data scope with the research questions put forward. The present dissertation addresses these challenges by examining how different measurement choices shape what can be validly inferred about digital media use and well-being in everyday life. To this end, I draw from a dataset collected within the ERC Starting Grant project DISCONNECT. Before presenting the empirical studies, I first present a literature review in which I elaborate on the paradigm shift towards behavioral data, to then introduce the central research questions guiding this dissertation.

Literature Review

In this brief literature review, I reflect on the paradigm shift from self-reported screen time behavior to behavioral data, to then discuss the challenges of theoretical interpretation, granularity and data scope.

Paradigm shift: From self-reported screen time to behavioral data

Accurately recalling the time spent on digital devices, apps or platforms is a cognitively demanding task: Digital interactions are frequent, distributed across multiple devices, and often fragmented throughout the day, making retrospective estimation inherently unreliable (Andrews et al., 2015; Deng et al., 2019; Tkaczyk et al., 2024). Digital devices are used for many different purposes, to engage with many different platforms, and often in short bursts, possibly even entailing many task switches (Deng et al., 2019; Jeong et al., 2020; Toth et al., 2025). Apart from this complexity, a normative[1] dimension to self-reporting digital behavior has been shown to be present as well. This can lead participants to both over- and underestimate digital behavior, based on what a “normal” amount of time should look like (Boyle et al., 2022; Johannes, Nguyen, et al., 2021).

Validation research confirms the discrepancies between self-reports and behavioral data. In a recent meta-analysis, Parry and colleagues (2021) found only moderate correlations between the two, concluding that self-reports often are unable to accurately represent actual behavior. These discrepancies are present in both directions, with studies such as Ohme and colleagues (2021) finding signs of clear underestimations, while studies such as Verbeij et al. (2021) showing overestimation of time spent on social media. The occurrence of these inaccuracies will hence impact the validity of any associations found between self-reported digital device use and well-being.

As a response, scholars have increasingly turned towards the collection of passively gathered behavioral data as a more objective alternative. These data are often considered the “gold(en) standard” (e.g., Goedhart et al., 2015; Leitgöb & Keusch, 2026). The use of these data, however, is not without problems or challenges. I will focus on three of these in the current literature review.

[1]For instance, Johannes et al. (2021, p. 2) state how “Broadly speaking, inaccuracy in the form of measurement error can be exclusively random or a mix of random and systematic error. For example, people systematically overestimate their physical activity (Klesges et al., 1990) and underestimate their smoking (Connor Gorber et al., 2009)—both are examples of social desirability, one important factor in shaping inaccurate self-reports of behavior (for an overview of other factors, see Schwarz & Oyserman, 2001).”

The first challenge: From behavioral objectivity to subjective experience

Digital trace data record when, what, and how long someone interacted with their device, but they do not reveal how these interactions were experienced. In research on well-being this distinction can be critical, as these outcomes-of-interest are often inherently subjective. Consider, for instance, that thirty minutes of social media use may be experienced by one person as a relaxing activity, and by another as procrastination that interferes with more important goals. In both cases, the behavioral record is identical, but the psychological implications may differ in ways that trace data alone cannot capture. Consequently, the theoretical interpretation of objective behavior as a representation of a particular psychological state is never straightforward, even when the underlying behavioral data are collected at a high temporal resolution.

The concept of digital wellbeing has recently been proposed to represent a dynamic system, where subjective experiences in relation to digital media should be seen as a function of person-, context-, and device-specific factors that dynamically interact over time (M. M. P. Vanden Abeele, 2021). Many of these factors, such as affective and cognitive appraisals of digital connectivity, or momentary motivational states, are inaccessible through behavioral observation alone, yet come with an underlying assumption that digital traces may give expression to them, as a 'digital phenotypic' manifestation of them. For instance, studies imply that 'availability pressure', defined as "a perceived obligation to respond quickly to incoming calls or messages and to regularly check the smartphone for potentially incoming notifications" might be represented by the behavior of frequently checking the phone for incoming notifications (e.g., 2026, p. 1). This underlying assumption suggests that the more fundamental question for digital well-being research is perhaps not whether behavior can be accurately measured, but whether these measures can be meaningfully translated into the subjective psychological states that they are meant to represent.

Recent theoretical work has begun to formalize this distinction. Wolfers (2024), for example, proposes a conceptual separation between objective media use and media use perceptions. Objective media use refers to externally observable behavior, e.g., duration and frequency, which can be captured through external agents such as logging apps. Media use perceptions, by contrast, refer to the mental model an individual has of their own media use, which Wolfers further differentiates into a knowledge and an evaluative component. This knowledge component encompasses descriptive perceptions and categorizations, such as one's own estimations of media use duration or classifying that media use as work or non-work, for example. The evaluative component consists of moral judgements and self-conscious emotions surrounding one's own media use: for example, evaluating it as excessive, labeling it as procrastination or recovery, and experiencing guilt in relation to this digital behavior. This distinction is consequential because, as Wolfers argues, evaluating media use negatively may be more likely to affect mental health than simply holding a descriptive perception of usage time. Earlier empirical work supports this claim, for example by showing that perceiving media use as procrastination versus recovery has a different association with well-being (Reinecke & Hofmann, 2016). These perceptions are not merely biased proxies for objective behavior; they represent distinct constructs with independent explanatory relevance (Sewall et al., 2020; Wolfers, 2024).

A growing body of empirical work provides support for the notion that subjective perceptions shape the relationship between digital behavior and well-being. Ernala et al. (2022) demonstrated that the association between time spent on the platform and subjective well-being was significantly moderated by users' beliefs about whether Facebook use was good or bad for themselves and for society: When participants believed that Facebook was good, time spent on Facebook was not significantly associated with well-being; when they believed it was bad, increased time spent was associated with lower well-being. This moderating effect was stronger when the time-spent measure was self-reported than when it was derived from server logs, suggesting that the subjective framing of one's behavior matters more for well-being than the behavior itself.

Work by Lee and Hancock (2024) further developed this line of work, coining the term 'social media mindsets', defined as "the core beliefs held by individuals that orient their expectations, behaviors, attributions, and goals regarding the role of social media in their lives" (p.2). They identified two dimensions: agency (perceived control over one's social media use) and valence (perceived effects as enhancing or harmful). These mindsets explained more variance in psychological well-being than conventional measures of social media use such as self-reported time spent or intensity. Lee and Hancock proposed two routes through which mindsets relate to well-being: an appraisal route, where mindsets moderate how individuals interpret their use (e.g., social media use predicted lower well-being for those with low-agency or negative-valence mindsets), and a behavioral route, where mindsets orient people toward different patterns of engagement (e.g., more agentic mindsets were associated with less passive use).

More recently, Parry and Coetzee (2025) incorporated both self-reported and logged social media use alongside social media mindsets. Their findings partially corroborated and partially challenged the earlier work. Agency mindsets were robustly associated with well-being across all four indicators, replicating Lee and Hancock's findings. However, valence mindsets were not consistently linked to well-being once logged behavior was controlled for. Importantly, when social media use was assessed through log data rather than self-reports, the associations between mindsets and well-being weakened substantially, and the evidence for the appraisal route became inconsistent. The authors concluded that while subjective experiences may be a better predictor of well-being than the logged amount of use, it may be perceived control over one's social media use, rather than general evaluative attitudes, that is most consequential (Parry & Coetzee, 2025).

This body of work yields several insights directly relevant for the use of digital trace data in well-being research. First, subjective perceptions of media use are not merely noisy reflections of objective behavior; they are likely separate constructs with their own explanatory power (Wolfers, 2024). Second, how users interpret and evaluate their digital behavior can moderate, or sometimes even overpower, the association between objective use and well-being (Ernala et al., 2022; Lee & Hancock, 2024). Third, not all subjective perceptions may be equally consequential. Perceptions of control appear to be more robustly linked to well-being than general evaluative judgements, and the strength of these associations depends on whether behavior is measured through self-report or through log data (Parry & Coetzee, 2025). This pattern aligns with Vanden Abeele's (2021) conceptualization of digital well-being as a dynamic system in which person-specific appraisals, rather than behavioral exposure per se, shape how media use becomes consequential for well-being. Both behavioral traces and subjective perceptions can hence be seen as complementary sources of evidence, each capturing a different facet of the media use experience (Wolfers, 2024). The question for digital well-being research is therefore not *whether* to use behavioral or subjective data, but *how* to meaningfully integrate both in order to better understand the mechanisms through which digital behavior relates to well-being.

If subjective experiences complement behavioral data, the methodological implication is that research designs should aim to capture both data streams in tandem, ideally at compatible temporal resolutions. This has motivated researchers to combine passively collected trace data with experience sampling methods, producing "linkage" designs that temporally align behavioral logs with momentary self-reports of subjective states and experiences (cf., Harari & Gosling, 2023; Klingelhoefer et al., 2025). Such designs allow researchers to examine not only whether behavior is associated with well-being, but through which subjective mechanisms associations emerge. In doing so, they operationalize complementarity between behavioral and subjective indicators that the theoretical and empirical work reviewed above calls for.

Emerging empirical work using such combined designs illustrates both the potential and current limits of this approach. The study by De Segovia et al. (de Segovia Vicente et al., 2024), to which I contributed, combined passively logged smartphone use with mobile momentary assessments to examine the downstream consequences of mindless scrolling, a specific form of social media consumption characterized by a reduced awareness and the absence of intentional goals. The findings revealed that longer periods of mindless scrolling were associated with increased guilt over smartphone use, a relationship that was partially mediated by the experience of goal conflict. Guilt accumulated over the course of a day and predicted lower end-of-day well-being. This psychological pathway – from behavior to goal conflict, through guilt, to diminished well-being – would be entirely invisible in behavioral data alone. The log data capture when and how long a person scrolled, but understanding whether and why that scrolling became consequential for well-being required access to the subjective mediators that connected the behavior to its outcomes.

Similarly, Elmer et al. (2025) combined smartphone log data with ESM to examine bidirectional within-person associations between smartphone usage and momentary well-being. Their results indicated that smartphone use in the hour before an ESM assessment, and particularly social media use, was associated with reduced affect valence and increased loneliness on the within-person level. Loneliness, in turn, predicted increased subsequent smartphone use, suggesting a bidirectional dynamic. However, the effects were consistently small, and the authors observed that individuals who generally scored higher on loneliness were more sensitive to the momentary effect of social media use. This between-person moderation, which aligns with susceptibility-based accounts of media effects (cf. Valkenburg et al., 2021), again shows the importance of capturing subjective states alongside behavioral indicators: the same amount of logged social media use related differently to well-being depending on the person's broader psychological profile.

The first aim of this dissertation is to contribute to the growing body of work focusing on the subjective mechanisms through which behavioral patterns become consequential for well-being (or perhaps rather fail to do so?). The theoretical and empirical work reviewed above establishes that digital trace data, while providing precise records of device interaction, lack the capacity to capture how users interpret and evaluate their own digital behavior – processes that appear to play a meaningful role in shaping well-being. Research designs that combine logged behavior with momentary subjective assessments can help address this gap, yet the field is still in the early stages of understanding which subjective mechanisms matter most, how they interact with specific behavioral patterns, and what this implies for the status of trace data as indicators of well-being. The first sub-research question of this dissertation therefore asks:

> **S-RQ1**: How do passively logged behavior indicators correspond with the cognitive-behavioral states they supposedly represent, and what does this imply for the status of log data as theoretically interpretable indicators of well-being?

The second challenge: Granularity and the operationalization of digital behavior

While including subjective experiences is essential for interpreting the psychosocial meaning of digital behavior, it does not by itself resolve the decision over which digital traces should be represented, and how they should be calculated. This shifts attention to an issue that perhaps precedes subjective appraisal: the level of granularity at which digital behavior is operationalized into ‘features’. The behavioral features selected for analysis serve as more than predictors; they define the space for possible associations with subjective states. Research that only operationalizes total duration, for instance, is unable to detect whether fragmented versus sustained use patterns of specific app categories relate differently to momentary experiences of stress or distraction.

Much of the existing research, whether using self-report or trace data, has relied on highly aggregated measures, most prominently total duration or frequency over a given time window (Parry et al., 2021). While these types of measures are perhaps more analytically convenient or intuitive, they treat digital media use as a homogenous concept (Kaye et al., 2020). This homogeneity assumption clashes with theoretical models that conceptualize digital well-being as a dynamic system in which person-, context-, and device-specific factors interact over time (M. M. P. Vanden Abeele, 2021). If well-being is shaped by the temporal patterns and contextual embedding of digital engagement, beyond its total volume, then aggregate indicators may be unable to capture the processes they are meant to represent. Recent empirical and theoretical work has increasingly questioned this assumption, and has indicated that, rather than the *amount* of time spent using digital technologies, *how* we spend or distribute that time is most relevant (Kaye et al., 2020; Meier, 2021).

Indeed, studies combining trace data with experience sampling show that momentary associations between well-being and digital media are often weak or inconsistent when operationalized in terms of total duration, yet become more meaningful when examined through the lens of more specific usage patterns. For instance, Rodman et al. (2024) provide empirical support for this claim, demonstrating that while overall smartphone use showed very weak or no associations with well-being, more granular features—such as engagement with distinct categories or platforms—were more predictive. Similarly, Elmer et al. (2025) found that total smartphone duration showed limited associations with well-being, while, when looking at specific app category analysis, social media appeared to be the main driver for these associations, with other categories largely showing no association. These findings lead to the conclusion that undifferentiated measures of screen time may systematically diminish observable associations with well-being outcomes.

Moreover, using digital traces allows researchers to go beyond coarse measures of “screen time”, enabling data collection at a level which can capture sequences or fluctuations of use (and non-use). This opens new possibilities to operationalize dimensions of digital behavior which would be very difficult to assess through the use of self-reports. Examples of such dimensions are fragmentation (Siebers et al., 2024), burstiness (Jo et al., 2012), or checking behavior or “glances” (Parry & Toth, 2025). These dimensions might be mapped more cleanly onto theoretically relevant behavior or psychological states, such as procrastination or distraction, than total duration or frequency alone (e.g., (Aalbers et al., 2021; Siebers et al., 2024). Many of these features are inherently dynamic: they capture not just the volume of behavior within a given period, but its temporal structure, i.e., how use is distributed over time. This temporal dimension aligns with theoretical calls for measurement approaches that match the timescale at which digital well-being processes unfold (Klingelhoefer et al., 2025; Vandenbosch et al., 2025), and its importance can be recognized in intensive longitudinal designs increasingly being adopted (e.g., Elmer et al., 2025; große Deters & Schoedel, 2024; Rodman et al., 2024).

These findings from extant research suggest that granularity should not just be reflected in data collection and afterwards flattened in highly aggregated measures, but rather be a core part of the operationalization of digital behavior. Aggregate indicators may obscure short-term processes, within-person variability and the broader context wherein this behavior takes place. Given that research embracing more granular operationalizations of screen time is still growing, whether this increased granularity translates into empirically stronger or more theoretically coherent associations with well-being, however, remains somewhat of an open question. The empirical chapters in this dissertation address this question by comparing how aggregate versus granular behavioral features – including category-specific duration, but also fragmentation and notification-based indicators – relate to momentary well-being states such as time pressure (Chapter 4) and mood (Chapter 5). Consequently, my second sub-research question asks:

> **S-RQ2**: How do granular and dynamic features of digitally logged behavior relate to momentary well-being outcomes in daily life?

The third challenge: Measurement scope and device coverage

Finally, a third limitation of much log-based research on digital media use and well-being concerns the limited scope of behavioral data collection. Despite methodological advances in passive sensing, digital trace data are most often collected from a single device, most commonly the smartphone, thereby implicitly treating this device as a proxy for someone's broader digital engagement (Parry et al., 2021). In measurement terms, this amounts to a form of construct underrepresentation: the operationalization of "digital media use" systematically excludes portions of the behavioral domain it aims to capture. This risks providing only a partial view of someone's everyday media behavior – a concern that extends beyond smartphone logging studies to other digital trace traditions, such as web tracking research, where the failure to capture data from all devices participants use to go online has been termed *tracking undercoverage* (Bosch et al., 2025).

A growing body of work has begun to acknowledge this limitation. For instance, a recent meta-analysis by Parry et al. (2021) shows that nearly all studies investigating log-based digital media use relied exclusively on smartphone data, with only a handful of studies incorporating passive sensing data on the use of computers, gaming devices or general internet use. Similarly, while studies have increasingly emphasized the importance of differentiating between different types of media use, differences in device use have often not been accounted for when examining well-being outcomes (e.g., Liu et al., 2024). Even work comparing the accuracy of self-reported social media use to log-based counterparts has explicitly acknowledged that their conclusions were limited to smartphone data alone, thereby excluding any usage on tablets, laptops or web-based versions of social media platforms (Verbeij et al., 2021).

Emerging evidence suggests that these are not merely theoretical concerns. In their web tracking study, Bosch et al. (2025) demonstrated that 74% of participants in an online panel had at least one device untracked, and that this undercoverage produced substantial biases in estimates of online media exposure, with the magnitude and even direction of bias varying depending on the statistic of interest. Similarly, Pankowska et al. (2025) showed, using hidden Markov models applied to Facebook tracking data and self-reports, that digital trace measures severely underestimated the frequency of Facebook use for approximately one-third of their sample, driven specifically by incomplete device coverage. While these studies focus on web tracking rather than smartphone logging, they provide a direct empirical warning: when the measurement instrument fails to capture behavior across all relevant devices, the resulting data can be substantially biased.

Consequences of this narrow measurement scope extend beyond underestimating overall levels of digital engagement. Research that only focuses on smartphone-based activity is unable to capture how cross-device patterns might shape associations with well-being. For example, receiving and reading a short email message on the smartphone can initiate a much longer working session on the computer. As we will demonstrate in Chapter 5, more than half of email screen time in our sample occurred on computers rather than smartphones, meaning that smartphone-only data would miss the majority of this particular category of use. Prior research already shows that being interrupted by, and needing to manage, overlapping roles and tasks can lead to negative well-being outcomes, such as exhaustion (Derks et al., 2021) and stress (Cornwell, 2013). In the above example, for instance, a smartphone-only study would therefore capture the *trigger* of such a sequence while remaining blind to the bulk of the *resulting behavior* and its potential well-being consequences.

When activities performed on certain devices are omitted, any observed association would therefore reflect only a subset of relevant behavior, leading to imprecise estimates or biased inferences of these effects. In such cases, these estimates could misrepresent the magnitude or even the direction of associations between digital media use and well-being. Critically, such bias need not be uniform or predictable. Bosch et al. (2025) demonstrated through simulations of device undercoverage that the direction and magnitude of bias depend on the statistic of interest: while simple count measures (e.g., total time online) are consistently underestimated when devices are missed, proportional or relational measures can be biased in either direction. In their data, undercoverage weakened certain correlations while inflating others. In the context of digital media use and well-being research, this means that single-device estimates could misrepresent not only the size but also the direction of associations.

More broadly, focusing on a single device constrains the kinds of theoretical questions that can be addressed. People's daily media repertoires are increasingly characterized by media multiplexity (Chan, 2015; Haythornthwaite, 2005), where communication takes place across multiple channels, platforms, and devices. Individuals do not interact with devices in isolation, but navigate a device-ecology, a system of interconnected screens across which attention, tasks and media engagement are distributed throughout the day. Capturing digital traces from computers alongside smartphones opens up the possibility of modelling cross-device dynamics, such as device switching, parallel use and multi-device workflows. As we will show in Chapter 5, such dynamics are far from rare: in our data participants switched between their smartphone and computer approximately 66 times per day, with nearly 90% of these switches also involving a change in the category of media activity. These patterns are invisible in single-device datasets, but could be central in emerging theories of media multitasking and attentional fragmentation (Abrahamse et al., 2025).

Ignoring computer-based log data should therefore not be seen as a minor omission, but rather a substantial limitation on what can validly be inferred about digital media use and its relationship with well-being. The evidence reviewed above, ranging from near-universal reliance on smartphone-only data in the literature (Parry et al., 2021), to the demonstrated biases that device-incomplete tracking introduces in media exposure estimates Pankowska et al. (2025), suggests that single-device measurement threatens both the content validity of digital media use constructs and the accuracy of their estimated associations with well-being outcomes. Hence, as a final sub-research question, this dissertation asks:

> **S-RQ3**: To what extent does reliance on single-device (smartphone) log data create systematic validity gaps in estimating digital media use and its association with well-being?

By situating this question at the intersection of measurement scope and theoretical inference, this dissertation aims to contribute to the methodological debate, but also to clarify the boundaries of what digital trace data can – and cannot – tell us about digital media use and well-being. Across the three sub-research questions developed above, a common thread emerges. Whether the challenge concerns the interpretative gap between behavioral indicators and cognitive-behavioral states (S-RQ1), the level of granularity at which digital behavior is operationalized (S-RQ2), or the scope of devices from which behavioral data are collected (S-RQ3), each reflects a different facet of the same underlying problem: the choices researchers make in turning raw digital traces into analyzable constructs shape, and possibly constrain, the validity of the conclusions that follow. Taken together, the main research question of this dissertation therefore asks:

> **RQ**: To what extent do the theoretical interpretation, granularity, and scope of digital trace data shape what can be validly inferred about digital media use and well-being?

Chapter Overview

This dissertation presents an inquiry into the above research questions, in the form of several empirical investigations. It includes three empirical chapters (Chapters 3-5) in which variations of digital trace measures are combined with self-reported well-being outcomes, contextual information, and subjective mechanisms linking digital behavior to experience. Each chapter addresses one or more sub-research questions, focusing respectively on psychological interpretation (Chapter 3), behavioral granularity (Chapter 4), and measurement scope (Chapter 5). Table 1.1 provides an overview of how each chapter contributes to each sub-research question.

In addition to the empirical chapters, Chapter 2 outlines the overarching methodological framework – including experience sampling design, digital trace data collection procedures, and analytic strategies – and situates the studies within the broader DISCONNECT project. Together, these chapters examine how digital trace data can be used, methodologically and theoretically, to study the relationship between digital media use and well-being.

Chapter 3 ("Connected Yet Cognitively Drained") investigates whether online vigilance promotes mental fatigue, especially under perceived availability pressure, and whether passively sensed smartphone behavior can serve as a digital proxy for this subjective state. By combining passively sensed behavioral features with self-reported measures of online vigilance and mental fatigue, this chapter tests whether behavioral indicators capture the same psychological phenomenon that self-reports do, and whether both pathways lead to comparable well-being outcomes. The findings reveal that behavioral smartphone features are only weakly associated with self-reported online vigilance, suggesting that trace data alone may not reliably represent the subjective experience thereof. This chapter thus primarily relates to S-RQ1, questioning the conditions under which passively logged measures can validly stand in for psychologically meaningful states. Like Chapter 4, this chapter computes its behavioral indicators exclusively from smartphone data, and the absence of computer-based traces is considered a scope limitation that Chapter 5 subsequently addresses.

Chapter 4 ("Always On, Always Rushed for Time") explores how momentary patterns of mobile communication, operationalized through four behavioral features (duration, frequency, fragmentation and notifications received) across four app categories (email, social media, chat, and work communication), relate to feeling rushed for time, both directly and indirectly through perceived task juggling load. As such, it shifts attention to the level of granularity at which digital behavior is operationalized. By demonstrating that granular, category-specific features reveal associations that aggregate measures obscure, this chapter primarily contributes to S-RQ2. Its finding that the link between behavioral patterns and time pressure is mediated by a subjective mechanism (perceived juggling load) also connects to S-RQ1, reinforcing that trace data require interpretative framing to become psychologically meaningful. As with Chapter 3, this study operates within a single-device (smartphone) measurement scope.

Chapter 5 ("Computers Matter") combines experience sampling with both passively logged smartphone and computer trace data, and investigates how the inclusion of computer-based and cross-device behavioral features – such as device switching, context switching, and overlapping use – changes both the descriptive estimates of screen time and the observed associations between digital media use and momentary and daily mood. This chapter directly addresses the measurement scope limitation that Chapters 3 and 4 acknowledged but could not resolve. It further demonstrates that smartphone-only measures substantially underestimate total digital engagement, that cross-device dynamics constitute a common yet previously underexamined dimension of everyday digital behavior, and that including computer-based and cross-device features meaningfully alters the observed associations between media use and affect. This chapter primarily contributes to S-RQ3, while its introduction of novel cross-device behavioral features also advances S-RQ2 by expanding the repertoire of granular indicators available for modelling digital media use.

Chapter 6 ("Making the Impossible Possible") complements the empirical work by addressing a practical barrier to the multi-device measurement strategies that our other chapters call for. In response to Apple's restrictive data access policies, which have historically prevented researchers from obtaining detailed behavioral data from iOS devices, this chapter presents a novel data donation procedure that leverages the built-in Screen Time synchronization feature across Apple devices to extract app-level usage data from iPhones, iPads, Apple Watches, and Macs linked to the same Apple ID. In addition, we provide an open-source parsing tool that converts the resulting system-level files into researcher-ready datasets. By offering a non-intrusive, privacy-sensitive, and technically accessible method for capturing iOS screen time data, this chapter directly contributes to S-RQ3, extending measurement scope beyond the Android-only focus of Chapters 3 and 4 (and much of the digital trace literature), and supporting multi-device data collection approaches that Chapter 5′s findings call for.

Chapter 7 (Discussion) synthesizes the findings of the preceding chapters and discusses their implications for the use of trace data in digital-wellbeing research, for broader debates within the field, and for the societal context in which screen time research is situated, before addressing the limitations of this dissertation.

Finally, beyond the chapters included in this dissertation, the collaborative nature of the DISCONNECT project has also allowed me to contribute to several related papers by colleagues within the project team.

Table 1.1: Overview of Dissertation Chapters and Research Questions

Chapter	Interpretation (S-RQ1)	Granularity (S-RQ2)	Scope (S-RQ3)
Ch. 3	*Central focus*	*Secondary*	*Acknowledged limitation*
Connected yet cognitively drained	Tests whether behavioral smartphone features can serve as proxies for self-reported online vigilance; finds weak correspondence, revealing the interpretative limits of trace data.	Constructs theory-driven behavioral features (monitoring sessions, notification responsiveness, reaction time) mapped to specific online vigilance dimensions.	Relies on smartphone data only; absence of computer-based traces limits measurement scope.
Ch. 4	*Secondary*	*Central focus*	*Acknowledged limitation*
Always on, always rushed for time	Shows that the link between smartphone use and time pressure operates through a subjective mechanism (perceived juggling load), reinforcing the need for interpretative framing.	Disaggregates smartphone use into category-specific (email, social, chat, work communication) and dynamic features (frequency, duration, fragmentation, notifications); shows these reveal associations that aggregate ones obscure.	Relies on smartphone data only; absence of computer-based traces limits measurement scope.
Ch. 5	*Secondary*	*Central focus*	*Central focus*
Computers matter	Shows that the same behavioral category (e.g., email) relates differently to mood, depending on device context, implying that device ecology shapes the psychological meaning of digital behavior.	Introduces novel cross-device behavioral features (device switching, context switching, overlapping use) that expand the repertoire of granular indicators beyond single-device metrics.	Shows that including computer log data substantially alters screen time estimates (+136%) and reshapes observed associations between media use and mood.
Ch. 6	*Not applicable*	*Not applicable*	*Central focus*
Making the impossible possible			Provides an accessible tool for collecting iOS device screen time data, enabling more inclusive multi-device data collection.

Methods

Introduction

This chapter documents the methodological foundations of the three empirical studies presented in this dissertation (Chapters 3-5). All three studies draw on the same overarching research design, combining intensive experience sampling and passively collected digital trace data from Android smartphones and, for a subset of participants, personal computers.

The chapter has two aims. The first is descriptive: to document the data collection procedures, processing pipelines, and analytic strategies shared across the empirical chapters. As the introduction chapter argued, the paradigm shift towards relying on digital trace data rather than self-reports intends to address some important limitations of self-reported media use (Ohme et al., 2021; Parry et al., 2021), but these trace data introduce their own measurement challenges. Digital traces do not speak for themselves, but require decisions about what is logged, how raw traces are cleaned and aggregated, and how resulting features are operationalized as indicators of theoretically relevant constructs (Olteanu et al., 2019; Parry & Klingelhoefer, 2025). These decisions are consequential, yet they are frequently relegated to appendices or supplementary materials, limiting the ability of researchers to evaluate and replicate data processing pipelines, and prompting calls for more transparent and detailed documentation of these pipelines (Baumgartner et al., 2023; Geyer et al., 2022; Parry & Klingelhoefer, 2025; Parry & Toth, 2025). This chapter aims to counter that tendency by offering a detailed overview of the entire trace-to-indicator workflow. Given that behavioral trace data become increasingly common in the social sciences, I thus hope this chapter can respond to a growing need for detailed accounts of the decision-making involved in transforming raw device logs into analytical variables, a process that is rarely reported in full (Parry & Toth, 2025).

The second purpose is reflexive: to make explicit how the many preprocessing and feature construction decisions shape the kind of behavioral indicators trace data can yield, and by extension, reflect on the theoretical claims they can support. This reflexivity is warranted because trace data are, for the most part, not designed with research in mind but are byproducts of platform and device architecture (van Atteveldt & Peng, 2018). Researchers must therefore make a series of choices: what data to retain, how to aggregate and label raw traces, and how to operationalize the resulting features as proxies for theoretically relevant constructs.

This concern is not unique to the present dissertation. Recent work has demonstrated that the number of defensible analytical choices available to researchers is vast. Using specification curve analysis, Orben & Przybylski (2019) showed how plausible "forking paths" in operationalizing technology use or well-being, or which co-variates to include, produce diverging effect estimates. Similarly, multiverse analyses, including one I contributed to (Murphy et al., 2025), show that preprocessing decisions directly shape the patterns of observed effects. The findings from any single analytical path may be accurate, but they are also contingent on decisions that are rarely present in standard reporting. This chapter responds to that problem not by implementing multiverse analyses, but by offering a detailed, fully documented account of the specific path taken in this dissertation: documenting the decisions, their rationale, and the constraints they impose. In doing so, the chapter also sets the stage for the empirical studies that follow, each of which illustrates how methodological choices alter what trace data can reveal.

To achieve the above purposes, the chapter is organized into five sections. The first section, titled **Data collection**, describes the study context. All studies were conducted within the DISCONNECT project, which took a particular approach to participant recruitment, and resulted in the capturing of three data sources: mobile experience sampling data, Android trace data, and computer trace data. The second section, titled **Data processing**, details the cleaning, event inference, and quality control steps applied to each trace data source. Next, in a third section titled **Combining Traces with Experience Sampling**, I will explain how behavioral logs were temporally aligned with ESM time windows to create the analytic datasets. Fourth, the **Feature Calculation** section presents the typology of behavioral features, from basic volume to cross-device metrics, computed across the empirical chapters. Finally, the **Analytic Strategy** section outlines the multilevel modeling approach, as well as the centering and temporal lagging procedures common in all three studies. Within each of these sections, I describe the decisions that I took and reflect on the challenges that came with them, and what downstream implications they have had for the empirical studies (and project).

Research Design

Data Collection

The data used for the studies in this dissertation were collected in wave 1 of the DISCONNECT project. Data collection for this wave took place between October and December of 2022 in Flanders, Belgium. Participant recruitment and scientific outreach occurred in collaboration with De Standaard, a major Flemish newspaper, enabling broad reach while introducing specific sampling considerations. Advertisements appeared on the newspaper's website and in print. Given that the newspaper targets an older, more educated audience, this impacted our sample composition: We recruited a slightly older and certainly more educated demographic compared to a representative sample. For instance, most were highly educated (93% had at least a bachelor's degree), with an average age of 38.83 ($SD = 11.72$).

Participants were incentivized through the offer of a personalized report, with insights into their own media usage patterns and associations with well-being, and free access to events and lectures related to the project. There were no monetary incentives for participation. This likely implies that participants were on average highly motivated to participate, perhaps because of their own experiences (and struggles) with digital well-being.

Participants took part in a two-week study period consisting of intensive mobile experience sampling while simultaneously collecting behavioral trace data from their Android and computer devices. In total, 3,065 participants expressed their interest in the study, of whom 1,315 eventually participated. Among these participants, 774 individuals provided Android trace data (after cleaning), and of this subsample, another 106 individuals also provided computer trace data.

Below, we provide a brief overview of each data source collected (mobile experience sampling, Android trace data, and computer trace data). Figure 2.1 provides a graphical overview of all data sources.

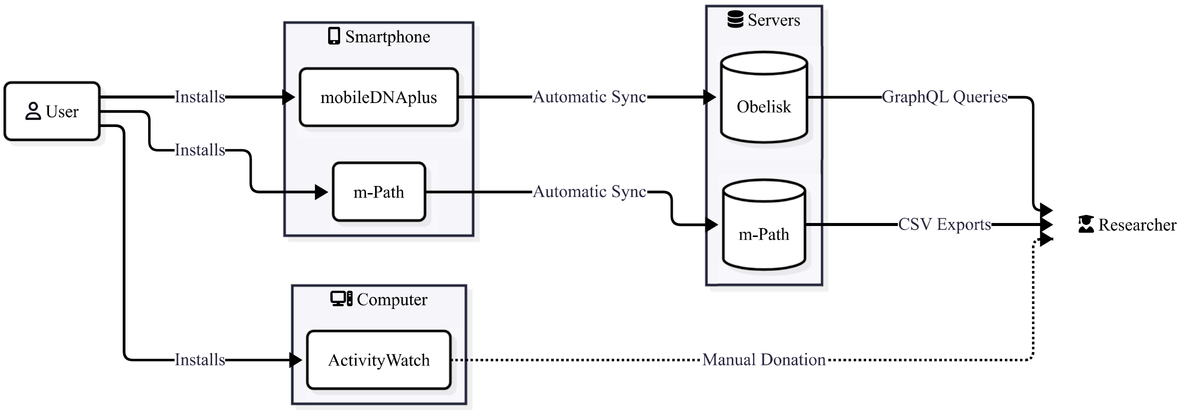


Figure 2.1: Overview of Data Sources and Architecture

2.2.a.a Mobile Experience Sampling

To carry out surveys and momentary assessments, we used m-Path[2], a tool developed as a KU Leuven spinoff, allowing researchers and practitioners to conduct momentary assessments. The tool works on both Android and iOS devices and enables asking short, in-the-moment questionnaires.

Each day, participants received six questionnaires distributed between 07:30 and 22:45, and this for a period of two weeks. Each questionnaire was scheduled at random within 90-minute time slots and remained active for 45 minutes with a reminder being sent after 30 minutes. This scheduling ensures temporal variability while maintaining sufficient separation between assessments (minimum 1h15m, maximum 4h15m).

Each first questionnaire of the day prefaced its questions with the stem “Since getting up…”, while all other questionnaires used “Since the last questionnaire…”. Participants were instructed to interpret this regardless of having completed this questionnaire or not. This retrospective framing was a deliberate choice with methodological consequences. Alternative ESM stems, such as “How … are you feeling right now?” or “In the past hour, …”, would have altered both the subjective construct being measured and the temporal alignment with trace data. A “right now” prompt captures momentary states, but creates ambiguity about which behavioral window it corresponds to (i.e., the preceding minutes, the hour before?). A fixed retrospective window (e.g., “in the past hour”) would standardize the recall period, but could also leave out important digital behavior that took place in the time before. Our chosen approach aimed to create a natural correspondence between the subjective report and behavioral trace window, as both relate to the same time span. However, it also means that participants are retrospectively evaluating periods of variable length, which may introduce systematic differences in recall quality, and shape which behavioral-outcome associations can be detected Klingelhoefer et al. (2025).

2.2.a.b Android Trace Data

Due to the strictness of Apple’s App Store policies, prohibiting apps designed to measure smartphone behavior, we were limited to collecting *Android* trace data, while iPhone users were asked to report daily on their Apple Screen Time metrics. Although m-Path now has access to mobile sensing, at the time of our project, this was still in development. As such, we employed a separate app to log mobile trace data.

Android participants were instructed to install ERC mobileDNAPlus, an app developed by the research group *IDLab* at the UGent Computer Science department. This app was loosely based

[2] https://m-path.io/

on a previous tool developed by external developers of digital product studio *In The Pocket* in collaboration with our own research group, *imec-mict-UGent*. IDLab also hosted the server infrastructure[3] platform, to which the app transmitted collected trace data[4].

Data transmission occurred at intervals determined by the app, and only when the participant's device was connected to Wi-Fi, to reduce any mobile data usage for the participant. The Obelisk platform provided us with data export and query capabilities, enabling extractions of raw trace logs for further processing.

The mobile tracing app collects three main data sources: App events, Screen events (or sessions), and Notification events, which were exported separately based on data source. As both mobileDNAPlus and its predecessor have unfortunately not been made open source, however, we cannot completely detail how raw smartphone trace data is collected. This reliance on closed-source tooling is a transparency limitation worth noting. Although we document the data structures and processing steps in detail below, the initial logging — what the app records, at what resolution, and under which conditions — remains partly opaque to us as researchers. This is a common challenge in behavioral sensing research, where scholars often depend on third-party tools whose internal logic is not fully accessible Geyer et al. (2022). It also means that the "raw" data we describe below are themselves ultimately already a product of certain design choices made by the developers, prior to any processing on our part.

2.2.a.c Computer Trace Data

As an optional step, participants could choose to install software that locally tracks computer trace data. We chose to use a third-party open-source software bundle, ActivityWatch[5], aimed at providing users with detailed, raw computer activity events from a Quantified Self mindset. By default, ActivityWatch enables two "watchers" or data streams: *aw-watcher-afk*, which checks if a user was active or *afk* (away from keyboard), based on mouse and keyboard activity, and *aw-watcher-window*, which tracks the active window, its title and URL. Our empirical chapter using computer trace data focused only on the *aw-watcher-window* data stream.

Out of data minimization principles, a custom version of the software was developed together with the founder of ActivityWatch. This version modified how the aw-watcher-window data stream handled event data before storage. For non-browser applications, the window title was removed, retaining only the executable name (e.g., "slack.exe"). For browser applications (e.g., Chrome, Firefox, Safari), the software applied a keyword matching logic to the window title: if the title contained a recognizable word – such as "facebook", "gmail", or "youtube" – it was replaced with a corresponding category label (e.g., "Facebook", "Email", "YouTube"). Browser traces without a matching keyword received the label "excluded". A similar matching procedure was also applied to the URL field. In all cases, both the raw window title and URL were removed after categorization, as these could contain sensitive private information and provided minimal added value for our intended analyses. The resulting category labels varied in granularity: some were platform-specific (e.g., "Facebook", "Instagram") while others related to broader functional categories (e.g., "News", "Work & Productivity"). These platform-specific labels were subsequently mapped to broader genre categories during data processing (see below). This

[3]https://idlab.ugent.be/resources/obelisk

[4]The original plan was to have mobileDNAPlus collect both trace data and experience sampling data, so that participants would only have to install one app rather than two. Unfortunately, IDLab was unable to meet its promise – while the questionnaire functionality was operable, the notifications functionality did not work, leading our team to make a last-minute switch to m-Path for collecting the ESM data.

[5]https://activitywatch.net/

custom version of the software can be found on GitHub: https://github.com/simonperneel/activitywatch_ERC.

Data from the *aw-watcher-window* data stream can be exported as JSON objects, containing a list of all events. Below we present an example of three window events. Each event contains a "data" object with the *app* name and *title*, a *timestamp* (ISO8601 formatted UTC), and a *duration* value (in seconds). An important distinction: duration represents time when a specific window was in the foreground. In other words, if multiple windows were active at the same time, only the "active window" duration was calculated. Ideally, this would be cross-checked with *aw-watcher-afk* data, to ensure all active, yet AFK durations are not included. Unfortunately, a large portion of participants did not export their AFK data, forcing us to rely only on active window durations. The absence of AFK data means that our computer duration metrics may overestimate active engagement: a participant who left their computer open on a document while grabbing a coffee would be seen as actively using that application. This limitation is particularly relevant for interpreting cross-device features (Chapter 5), where simultaneous smartphone and computer activity may partly reflect a computer session left passively open rather than genuine multi-tasking.

As ActivityWatch keeps the data stored locally, participants exported and transferred their data file(s) through a secured file-transfer service. An example of the window data can be found below.


```
{
    "data": {
        "app": "EXCEL.EXE",
        "title": null
    },
    "duration": 1.099,
    "timestamp": "2022-11-02T23:14:24.710000+00:00"
},
{
    "data": {
        "app": "chrome.exe",
        "title": "Email"
    },
    "duration": 19.443,
    "timestamp": "2022-11-02T23:23:02.765000+00:00"
},
{
    "data": {
        "app": "chrome.exe",
        "title": "excluded"
    },
    "duration": 77.169,
    "timestamp": "2022-11-02T23:21:44.525000+00:00"
}
```


Example data of our custom-version of ActivityWatch:
Three window events, of which the second shows that the software found a match based on URL or title and remapped the "title" value to "Email". The third event presents a browser event where no mapping can be made.

2.2.a.d iOS Trace Data

As mentioned above, due to Apple's strict policies, the development of a parallel iOS trace data app was not possible during our data collection phase. Chapter 6, however, presents a new method developed after our main data collection period ended which *does* allow for granular screen time tracing on Apple devices. Please refer to the chapter for further details.

Data Processing

In this section, we describe all data processing steps that went into transforming raw trace data into processed datasets suitable for linkage with experience sampling data.

2.2.b.a Processing Android Trace Data

2.2.b.a.a Screen Data

Raw Screen data is presented in Table 2.1. Each row contains a timestamped screen event, which has four distinct options (see the "value" column): "Off", "On", "Locked", and "Unlocked". The "metric" column labels the type of data source, the "source" column contains the unique participant identifier, and the "tsReceived" timestamp column records when the server received a (batch) of data from the smartphone.

Table 2.1: Screen data (raw).

Timestamp	Metric	Value	Source ID
15:52:09.281	smartphone.screen	Off	b8af69c1
15:52:09.281	smartphone.screen	Locked	b8af69c1
15:52:10.130	smartphone.screen	On	b8af69c1
15:52:15.066	smartphone.screen	Unlocked	b8af69c1

Of most interest to us was knowing when screens were active and inactive. The act of locking or unlocking the device is mostly very close in time to the related screen "On" or "Off" events. However, not all these events are. For example, imagine someone turning on the screen to check a notification (preview) and then turning the screen off again. These logs would only include the "On" and "Off" values but exclude "Unlocked" and "Locked" events. We chose, however, to parse the Screen data into the format shown in Table 2.2, with each row containing an "On" and "Off" event and timestamp.

Step 1: Data processing started by sorting the data chronologically within participants based on the timestamp of events.

Step 2: Next, screen-state events were filtered to retain only "On" and "Off" events, discarding "Locked" and "Unlocked" states.

Step 3: Because each row recorded only the timestamp of its screen activity, we needed to determine when that event ended, i.e., for each "On" event, we needed to find a corresponding "Off" event. We inferred this by treating the start timestamp of the subsequent event (within the same participant) as the stop timestamp of the current event. Concretely, for each row, we copied the start timestamp and screen sate from the following row into two new columns – labeled *session_stop* and *shift_screen,* respectively– so that each row now contained both its own start information and the start information of the next event (an operation sometimes referred to as *shifting* or *lagging* in data processing).

Step 4: Finally, we noticed sequences where consecutive events shared the same state (e.g., two successive "On" events, likely indicating a missed "Off" event). These sequences were removed, so that only rows were retained where an "On" event was immediately followed by an "Off" event.

The above procedure has implications: By collapsing the four screen states into On/Off pairs, we prioritized a clean metric of screen-active duration over a more nuanced account of lock/unlock behavior. This means, for instance, that brief screen activations without unlocking, such as glancing at a notification preview, are counted as screen-on time but are not distinguished from fully unlocked, interactive sessions. An alternative approach could have preserved the Locked/Unlocked distinction, enabling separate metrics for "glance" versus "engaged" episodes. We chose the simpler representation, however, for three reasons: our primary interest was total active screen time as an input to further feature calculation; the On/Off structure facilitates annotation of App event logs with corresponding Screen sessions (see Application Data processing below); and the locked vs unlocked distinction is partially captured by our monitoring feature used in Chapter 3. However, this choice means that our screen duration metrics may slightly overestimate active engagement.

Table 2.2: Screen data (parsed).

uuid	session_start	screen	session_stop	shift_screen
b8af69c1	09:07:32.043	On	09:07:32.518	Off
b8af69c1	09:08:18.801	On	09:14:29.467	Off
b8af69c1	09:18:21.543	On	09:20:29.568	Off
b8af69c1	09:39:57.207	On	09:42:20.572	Off

2.2.b.a.b Application Data

Application event data contain a starting *timestamp* and a *packagename* (i.e.: the internal application name used in Android, under the "value" column). Table 2.3 shows an example, while Table 2.4 presents a minimized version with only relevant columns for further processing.

Table 2.3: Application usage data (raw).

timestamp	metric	value	source	tsReceived
1668502963450	smartphone.application	be.argenta.bankieren	b8af69c1	1668730009627
1668502963944	smartphone.application	com.android.systemui	b8af69c1	1668730009627
1668502969162	smartphone.application	com.google.android.input...	b8af69c1	1668730009628
1668502972788	smartphone.application	be.argenta.bankieren	b8af69c1	1668730009628
1668502973820	smartphone.application	com.google.android.input...	b8af69c1	1668730009628

In its unprocessed state, it would be impossible to calculate accurate screen time metrics, as there are only events *starting*, never *stopping*. This shows that preliminary data cleaning needs to be considered to correctly account for application screen time.

The example application data in Table 2.4 starts with a banking app (*be.argenta.bankieren*) being opened, followed by a system process (*com.android.systemui*) and the Google keyboard "app" launching (*com.google.android.inputmethod.latin*). Below, we present the five-step process of processing application data.

Step 1: We chose to remove keyboard application events, to attribute all time spent using the keyboard within an app as part of that app's screen time. This decision treats keyboard use as part of the parent application rather than a distinct behavioral event. An alternative would have been to retain keyboard events and attribute them separately, which could allow for distinguishing between passive consumption (e.g., scrolling) and active production (e.g., typing a message). However, as keyboard events in our data lacked both metadata about the parent app context in which they occurred and stop timestamps, accurate attribution would have been unreliable. For example, consider a participant scrolling through a social media app who receives an incoming message notification. The participant can reply directly via the notification tray without ever foregrounding the messaging app. In our data, the resulting keyboard event would carry no metadata indicating which app context triggered it, making it impossible to reliably attribute the typing activity to either the social media or messaging app. We therefore opted for the more conservative approach of absorbing keyboard time into the preceding application.

Step 2: Next, we needed to infer a stop event for each app, indicating that the app was now no longer active in the foreground. We did this with four additional steps. First, we sorted the data frame on *timestamp* and *participant*, then shifted each *startTime* and *app* to its previous row, when matching the same *uuid*. Table 2.5 shows this data frame.

Second, we identified the boundaries between different apps by comparing package names of consecutive events. When two successive events involved different package names, this transition marked a boundary. When consecutive events shared the same package name, they were treated as belonging to the same uninterrupted sequence of use for that app. We assigned a *run identifier* – a unique label for each such connected sequence – so that, for example, three consecutive events for *be.argenta.bankieren* would all receive the same run identifier, distinguishing them from a later, separate sequence of events for the same app.

Third, these data were aggregated at the level of the *participant* and *run identifier*. For each run we then extracted:

- The app associated with the run (i.e. the package of the first event)
- The start time of the run (i.e. start time of the first event)
- The stop time of the run (i.e. end time inferred from the last shifted start time)
- The app associated with the final shift event (as a consistency check)

Fourth, we retained the columns required for further analysis and sorted the dataframe chronologically within participants. An example of the resulting dataframe is shown in Table 2.6. Here, we can see that the consecutive rows where the *be.argenta.bankieren* app was used are now *flattened* into a single application event.

Step 3: We annotated Application data with Screen data. This allowed us to define stop times of Application events which ended due to the screen turning off, instead of a new app being opened. For each Application event, we looked backwards to identify the closest "Screen On" event and forward to the closest "Screen Off" event. These values were then added to each row of the data frame. See Table 2.7 for an example and refer to Table 2.2 for the relevant Screen events which were added.

Step 4: With this data frame we can finally calculate the correct Application stop times and, following, their duration. We did so by calculating an effective stop time value, which takes the earliest value between the Application stop time and the annotated Screen end time. Duration was then calculated as the difference between the start time and the effective stop time.

Step 5: As a final step in preparing the Application data frame, we included meta-data of more than 4.000 packages, including app name (e.g., "Instagram" for package "com.instagram.android") and genre, which we obtained through the Google Play Store (e.g., "Social" for com.instagram.android), as well as a custom genre label we manually added for apps missing this information or when the genre provided by the Play Store was deemed ill-fitting. For instance, the Google Play Store genre labels do not always align with how apps are used in practice: Google Chrome is for example labeled as "Communication" in the Play Store, which we changed into "Browser", as we preferred to only label chat and messaging apps as "Communication" (e.g., WhatsApp, Signal). Our custom labels attempted to correct the most prominent mismatches and better align categories with those relevant to our research questions. Any categorization scheme inevitably simplifies the multifunctional nature of most apps, however (J. A. Hall, 2023; M. M. P. Vanden Abeele et al., 2025). For instance, an app classified as "Social Media" may also serve as a primary channel for private messaging (J. A. Hall, 2023), and platform-level genre labels, such as those provided by app stores, are themselves unstable and inconsistently applied across operating systems (M. M. P. Vanden Abeele et al., 2025). This is particularly relevant for our categorical features (see Feature Calculation), where the theoretical meaning of, for instance, "social media duration" clearly depends on which apps were included in that category.

Table 2.4: Application data (parsed #1).

uuid	startTime	app
b8af69c1	09:02:43.450	be.argenta.bankieren
b8af69c1	09:02:43.944	com.android.systemui
b8af69c1	09:02:49.162	com.google.android.input...
b8af69c1	09:02:52.788	be.argenta.bankieren
b8af69c1	09:02:53.820	com.google.android.input...

Table 2.5: Application data (no keyboard events, shifted).

uuid	startTime	app	shift_start	shift_app
b8af69c1	09:02:43.450	be.argenta.bankieren	09:02:43.944	com.android.systemui
b8af69c1	09:02:43.944	com.android.systemui	09:02:52.788	be.argenta.bankieren
b8af69c1	09:02:52.788	be.argenta.bankieren	09:02:55.392	be.argenta.bankieren
...	...	...	...	...
b8af69c1	09:04:28.009	be.argenta.bankieren	09:04:31.575	com.android.launcher3
b8af69c1	09:04:31.575	com.android.launcher3	09:04:31.707	com.google.android.google...

Rows omitted above (…) are subsequent application events of be.argenta.bankieren.

Table 2.6: Application data (continuous app sequences).

uuid	app	start	stop	shift_app
b8af69c1	be.argenta.bankieren	09:02:43.450	09:02:43.944	com.android.systemui
b8af69c1	com.android.systemui	09:02:43.944	09:02:52.788	be.argenta.bankieren
b8af69c1	be.argenta.bankieren	09:02:52.788	09:04:31.575	com.android.launcher3
b8af69c1	com.android.launcher3	09:04:31.575	09:04:31.707	com.google.android.google...

Table 2.7: Application data (annotated with Screen events).

uuid	app	start	stop	session_start	session_stop
b8af69c1	be.argenta.bankieren	09:02:16.199	09:02:37.009	null	09:07:32.518
b8af69c1	com.android.systemui	09:02:37.009	09:02:43.450	null	09:07:32.518
b8af69c1	be.argenta.bankieren	09:02:43.450	09:02:43.944	null	09:07:32.518
b8af69c1	com.android.systemui	09:02:43.944	09:02:52.788	null	09:07:32.518
b8af69c1	be.argenta.bankieren	09:02:52.788	09:04:31.575	null	09:07:32.518

2.2.b.a.c Notification Data

Notification events are the third and final data stream captured by the Android application. Each row contains a participant identifier ("uuid"), a corresponding app ("app"), a timestamp signifying when the notification action "POSTED" took place ("postedOn"). Next, the "action" column shows whether a new notification was created or an existing one updated, i.e. "POSTED", or whether a notification was removed and no longer visible, i.e. "REMOVED". These removals can occur through different processes; the integers in the "reasonRemoved"[6] column clarify these. The "clicked" column shows when a "REMOVED" action was triggered due to the user clicking on the related notification. Finally, the "id" column contains a notification identifier, while "time" provides the timestamp of the current action. An example of notification data can be found in Table 2.8.

Table 2.8: Notification data (raw, with irrelevant columns reduced).

uuid	app	postedOn	action	reasonRemoved	clicked	id	time
c507d674	com.whatsapp	11:03:43.360	POSTED	null	null	1	11:03:51.984
c507d674	com.whatsapp	11:01:32.288	REMOVED	6	false	1	11:03:52.249

Without processing, notification traces showed massively inflated numbers, as these included both background (i.e. "system"-based) and foreground (i.e. those a user can view/interact with) notification events. Again, these data required multiple cleaning steps to become usable for metrics and analysis.

Step 1: Data cleaning started by dropping all duplicate rows and dropping rows where a notification event had the same *uuid*, *app*, *action*, and a *timestamp* at the same second.

Step 2: In addition, we removed all notification events for three specific packages: *android*, *com.android.systemui*, *com.google.android.apps.maps*. The first two packages appeared to be system-only notifications which a user does not interact with; the third had highly inflated numbers, as, for instance when navigating a route with Google Maps, each direction update is logged as a separate notification event. These first two steps already removed 61.9% of all notification events in the dataset (to N = 4,123,138). The case of Google Maps, however, illustrated a broader challenge: notification data conflate user-facing notifications with system-level events that are logged identically. Our removal decisions were based on manual inspection of the most frequent notification sources, but less common system-generated notifications from other apps may still be present in the dataset. This is an inherent limitation of working with

[6] A full list of these values and their meaning can be found in the Android Developer documentation, under the Constants section (see: REASON_*) https://developer.android.com/reference/android/service/notification/NotificationListenerService

device-level notification logs, which do not natively distinguish between notifications a user perceives and those generated in the background.

Step 3: Next, we refined the dataset by looking at the *id* column. The *id* column does not contain unique values but rather reuses identifiers over time. Many notifications appeared to be related to the same id and action, however. We started by sorting the data frame by participant and timestamp, and creating shifted versions of key variables (uuid, app, action, and id) backwards, i.e. the attributes of the next notification in the sequence. By comparing these, we looked for changes in the sequences. Whenever any of these attributes changed, we considered this a new notification event. A run identifier was added to keep track of these changes and allowed for "duplicates" to later collapse onto themselves. We did this by taking the first occurrence of stable attributes within these groups (e.g., the app, action and notification id) and the last occurrence of time-varying or shifted attributes (e.g., timestamp and app of the subsequent event). This further deduplicated the dataset by 59% (to N = 1,709,011).

Step 4: As a last step, we again added category (or genre) metadata to allow for analyses at the categorical level. An example of the cleaned notification data frame is shown in Table 2.9.

Table 2.9: Notification data (cleaned).

uuid	time	app	action	reasonRemoved	clicked	id	custom_genre
c507d674	17:57:52	com.whatsapp	POSTED	null	null	1	chat
c507d674	18:03:41	com.whatsapp	REMOVED	8	false	1	chat
c507d674	18:12:46	be.smartschool.mobile	REMOVED	1	true	451	education
c507d674	18:28:58	com.snapchat.android	POSTED	null	true	1943	chat

2.2.b.b Processing Computer Trace Data

The processing pipeline began by merging individual JSON files into a single data frame. Donated data files were organized in separate per-participant folders, named by each participant's unique identifier, allowing us to annotate each participant's data accordingly during import.

Next, we integrated the browser-level category labels produced at the collection stage (see above) with a broader application-level categorization. We generated a frequency-ranked overview of all programs and browser categories appearing in the data. For the 221 most frequently occurring entries – which together accounted for 97% of all traces – we manually assigned a standardized display *name* and genre *category* (e.g., OUTLOOK.EXE was given the name "Outlook" and category "Email"). For browser traces that had already received platform-specific labels from the custom software (e.g., "Facebook", "Instagram"), this step assigned broader genre categories (e.g., "Social Media"). Browser traces that had not matched any keyword during collection, previously labeled "excluded", were relabeled as "Browser – Other". Traces from applications outside the top 221 were labeled "Other". These were added to all applicable traces, while traces without a matching category were labeled as "Other". Table 2.10 illustrates the final structure of the processed computer trace data.

Table 2.10: Computer Trace Data (processed).

uuid	start	stop	duration	app	category	name
052536e1	14:49:36.737	14:49:39.007	2.27	slack.exe	work communication	Slack

052536e1	14:49:40.149	14:49:42.419	2.27	Spotify.exe	music&audio	Spotify
052536e1	14:49:43.525	14:49:55.765	12.24	slack.exe	work communication	Slack
052536e1	14:56:28.075	14:56:29.156	1.081	Signal.exe	chat	Signal
052536e1	15:19:15.720	15:19:35.53	19.81	slack.exe	work communication	Slack

2.2.b.c Quality Control and Participant Exclusion

After calculating application durations, we looked at cleaning outliers and removing irrelevant Application events which should not count towards someone's screen time. Before cleaning, the dataset contained 9,953,496 events. We identified duration outliers as Application events exceeding 240 minutes and dropped these from the data frame (N=5,617 or 0.056%%). This threshold of a single uninterrupted 240-minute app event is unlikely to reflect genuine continuous use, and more plausibly indicates a logging or data processing error, or an app running in the background. However, we acknowledge that any fixed threshold involves a trade-off. A lower threshold would be more conservative but risks excluding genuine long sessions, for instance a participant watching more than 4 hours of feature film through a streaming app. A higher threshold would retain more data but introduce more noise from logging artefacts. We examined the distribution of single app event durations and found that 240 minutes represented a clear inflection point beyond which events became extremely sparse and were often associated with known system apps. As we only collected start times of Application events, and had to rely on imputed stop times, we felt we struck the right balance with this threshold. Researchers applying similar pipelines should be aware that this threshold, like all cleaning decisions, shapes the resulting distributions of duration-based features and should be calibrated to their specific research context.

Next, we removed all rows for which we had no *stop* or *session_start* timestamps, removing an additional 60,393 rows (or 0.61%).

In addition, we removed Application events from apps we deemed irrelevant when calculating screen time metrics, such as: Always On Display (AOD) apps, fingerprint sensor apps, and apps without package names (N=1,552,835 or 15.71%). As users do not really engage with these applications, we felt these could be omitted. A full list of these apps is included on Chapter 3's OSF page: https://osf.io/bd5jt/. Finally, we looked at duration outliers for system-related applications, such as "android" and "com.android.systemui", calculated specific thresholds for each and dropped them from the dataset (N=114,973 or 1.38%).

Certain Android devices (such as Huawei, Xiaomi, and Samsung models) had operating system restrictions which limited ongoing data logging. We implemented a monitoring dashboard to detect data loss patterns as quickly as possible and provided device-specific instructions to restore tracking permissions (see https://dontkillmyapp.com/). Nonetheless, when calculating day-level trace data metrics, participants were removed from the dataset if they had fewer than seven days of data. Descriptives at the day-level (e.g. average daily screen time) were calculated by first removing everyone's first and last days of log data, to account for "incomplete" days.

Finally, the ESM dataset was cleaned using two criteria that were pre-registered: only participants with at least eight completed questionnaires were retained, and the completion time of a questionnaire had to be thirty seconds or longer.

Combining Traces with Experience Sampling

After initial trace data processing, the next step in the processing pipeline entailed combining trace data and experience sampling data. Our final datasets for analysis were focused at the level of each experience sampling beep, allowing us to model momentary associations. In other words, each trace needed to be labeled with the correct ESM time window to allow for aggregation at the same time span.

Earlier, we explained that our ESM schedule contained six beeps daily for two weeks. An important consequence of using variable-length ESM windows is that the same absolute amount of screen time (e.g., 30 minutes) represents a different proportion of the available window depending on whether that window was 75 minutes or 4 hours long. While our within-person centering and standardization procedures (see Analytic Strategy) partially account for this variability, researchers should be aware that raw feature values are not directly comparable across windows of different lengths. We did not normalize features by window length, as doing so would alter the interpretative meaning of the variable (e.g., from "how much was used" to "what proportion of available time was used"), which may not always align with the theoretical construct of interest. We present Table 2.11, Table 2.12, and Table 2.13 as an example of one participant's ESM windows, trace data (unlabeled) and trace data (labeled with ESM windows) to make clear how we assign each app episode with a matching questionnaire window. This labeled dataset then provides the basis for aggregating trace data into beep-level indicators (such as total duration and frequency of app and app categories) used in subsequent analyses.

An important prerequisite before combining experience sampling and trace datasets was ensuring temporal compatibility. Both datasets needed to use consistent timestamp formats and time zone representations, and participants had to be linkable across datasets through shared unique identifiers. In our case, aligning timestamps required particular vigilance regarding Daylight Saving Time transitions that occurred during the data collection period, as well as any instances of participants crossing time zones. We detected misalignment by cross-referencing data sources, for example, identifying cases where a participant had completed an ESM questionnaire at a given time, yet no corresponding m-Path app event appeared in the trace log at that timestamp. Resolving such discrepancies requires systematic inspection of individual participant timelines. We draw attention to this step because timestamp alignment in multi-source data is often more labor-intensive than anticipated, and errors at this stage propagate silently into all downstream feature calculations and momentary modelling.

Table 2.11: Example ESM windows.

uuid	date	question	time_wakeup	time_sent	time_start
052536e1	17/11/2022	Q1	07:20	08:00	08:03
052536e1	17/11/2022	Q2		10:00	
052536e1	17/11/2022	Q3		12:00	12:01

Table 2.12: Example Trace Data (to be labeled).

uuid	start	stop	app
052536e1	07:45	07:55	Instagram
052536e1	08:10	08:20	WhatsApp
052536e1	09:10	09:12	Gmail
052536e1	10:10	10:25	Chrome
052536e1	11:30	11:40	Chrome

Table 2.13: Example Trace Data (labeled with ESM windows).

uuid	app	start	stop	question	window_start	window_end
052536e1	Instagram	07:45	07:55	Q1	07:20 (wake-up)	08:03 (answered start)
052536e1	WhatsApp	08:10	08:20	Q2	08:00 (prev sent)	10:00 (sent; missed)
052536e1	Gmail	09:10	09:12	Q2	08:00	10:00
052536e1	Chrome	10:10	10:25	Q3	10:00 (prev sent)	12:01 (answered start)
052536e1	Chrome	11:30	11:40	Q3	10:00	12:01

Feature Calculation

All features calculated across the three empirical chapters were derived from the combined ESM-trace datasets such as described above, where each behavioral trace is aggregated to its corresponding ESM time window. This temporal alignment enabled modeling of momentary associations between digital behavior and subjective well-being states. The features can be organized into a typology that reflects increasing levels of theoretical specificity and behavioral complexity.

At the most basic level, **volume** features quantify the intensity of digital engagement through simple aggregation of trace events within each ESM time window. These form the foundation upon which more complex indicators are built. *Duration* sums the total time spent within specific applications (from app and/or computer data). *Frequency* counts the number of discrete events: app launches (from app and/or computer data), screen activations (from screen data), or notifications received (from notification data). While volume features alone reveal little about the nature of the engagement with an app, they provide essential baseline metrics. These features capture total exposure, which theories of media effects have long positioned as a predictor of (well-being) outcomes (although recent work questions this assumption, e.g., (Brinberg et al., 2023; große Deters & Schoedel, 2024; Kaye et al., 2020; Schenkel et al., 2024)).

The same volume features become more meaningful when disaggregated by **app category**. Rather than treating all screen time as equivalent, we categorized screen time into theoretically meaningful groups: *email*, *work communication*, *social media*, *chat*, and *total* (aggregated across all categories). This disaggregation aligns with calls to move beyond screen time as a monolithic construct (Kaye et al., 2020; Meier & Reinecke, 2021). Categorical features allow an increased alignment between behavioral indicators and theoretical expectations. Even though the trace data do not provide insight into *content*, they acknowledge that *what* people do digitally matters next to *how much* they do it. For example, in Chapter 4, when examining time pressure, the theoretical expectation differs by category. Email and work communication features are hypothesized to increase feelings of being rushed, as they can signal increased work demands, while social media might have more neutral or even restorative associations.

While volume features look at "*how much?*", and categorical features answer "*what kind?*", fragmentation features address "*how is this behavior distributed in time?*". These features capture the **temporal structure** of digital engagement, revealing whether screen time accumulated across multiple short bursts or rather through sustained sessions. For example, two participants might each use smartphone email apps for 30 minutes within a 3-hour ESM window. Participant A could have one 30-minute session (low fragmentation), while Participant B could have fifteen 2-minute sessions (high fragmentation). Despite identical duration, the dynamics of this behavior are drastically different. These dynamics could potentially impact related psychological outcomes, as highly fragmented behavior could, for example, signal more frequent task-switching, attentional disruption or habitual checking (Siebers et al., 2024; Van Gaeveren et al., 2025).

In other words, fragmentation quantifies the dispersion of duration and frequency across an ESM time window. Low values indicate concentrated use in fewer, longer sessions. High values indicate that the same total duration or frequency is dispersed across many brief episodes. Several operationalizations of fragmentation have been proposed in the literature. Alexander et al. (2010; 2011) developed multidimensional indicators of spatial and temporal activity fragmentation in the context of transport research, while Pan et al. (2019) proposed temporal stability parameters for smartphone use patterns, primarily aimed at identifying compulsive use. More recently, große Deters and Schoedel (2024) developed fragmentation indicators specifically for smartphone sensing data, capturing the distribution of usage and non-usage episodes. We adopted the operationalization proposed by Siebers et al. (2024), as their metric was designed for ESM-linked smartphone trace data and aligns most directly with our time-windowed feature calculation approach.

Next, some features were explicitly designed to operationalize **theoretically relevant** constructs. These aimed to serve as proxies for certain behavioral states or processes, based on the theoretical reasoning of what observable behavior might represent these constructs. In Chapter 3, we aimed to operationalize these proxies for the construct of online vigilance, defined as a state of heightened awareness, monitoring, and reactibility to digital communication (Reinecke et al., 2018). Concretely, we examined *monitoring behavior* by capturing the number of times a participant unlocked their phone for 5 seconds or less, without responding to a notification. This behavior reflects self-initiated "checking-in" on one's digital environment. Next, *responsiveness* was calculated as the percentage of notifications clicked relative to those received. High responsiveness suggests heightened reactibility, as users are acutely aware of, and interact with, incoming notifications. Finally, we operationalized *reaction time* as the average time (in seconds) between receiving a notification and clicking it. Fast reaction times could also indicate increased reactibility. These three features represent an attempt to map observable smartphone behavior onto psychological constructs, but as Chapter 3 will show, logged behavior may not always serve as a valid indicator of these subjective experiences.

All features described above can be calculated from single-device trace data. In Chapter 5, we introduce **cross-device** features, where smartphone and computer data are integrated to try and capture how people coordinate digital activities across devices. We attempted to go beyond additive logic, e.g. total screen time is the sum of smartphone and computer screen time, to also reveal patterns which can only be observed when multiple devices are tracked simultaneously. *Device switching* counts the number of times a participant alternated between devices. A switch occurs when either a participant stops using Device A and immediately switches to Device B, or, during active use of Device A, Device B is used. Next, *context switching* extends the previous feature to identify when a device transition also involves a change in activity category. For example, moving from email computer use to smartphone social media use would be seen as a context switch. Finally, *overlapping device use* measures the proportion of screen time where both devices were active simultaneously, calculated separately per device. These cross-device features assume that temporal co-occurrence of device activity reflects meaningful behavioral coordination. However, as noted above, the absence of AFK data for computers means that some instances of "overlapping use" may reflect a passively open computer, rather than active multi-device engagement. Similarly, device switching is identified through temporal proximity of events across devices, but the threshold for what counts as "immediate" involves a researcher-defined parameter. These features should therefore be understood as an approximation of cross-device dynamics rather than precise measurements.

The final level of our typology recognizes that even basic behavioral features gain psychological meaning through **subjective processes** or the **context** in which they occur. Rather than treating behavioral features as standalone predictors, here, we model interactions or mediations with subjective experiences. For example, Chapter 4 discusses the relationship between multiple communication features and feeling rushed, where we hypothesized that this relationship would operate through perceived juggling load (the experience of having to juggle multiple tasks). For example, the same amount of email notifications could feel overwhelming when juggling work and personal roles but be manageable when focused purely on the work domain. This approach acknowledges that behavioral features may be most informative when combined with subjective reports that reveal how digital interactions are experienced.

Table 2.14 contains an overview of all features used throughout this dissertation.

Table 2.14: Feature levels of digital trace data.

Feature Level	Examples	Information	Chapters
Volume	Duration, Frequency	Basic exposure metrics	3, 4, 5
Categorical	Email duration, Email frequency	Content-specific patterns	4, 5
Temporal structure	Fragmentation	Distribution of behavior in time	4, 5
Theory-aligned	Monitoring, Responsiveness, Reactivity	Behavioral proxies for psychological constructs	3
Cross-device	Device switching, Context switching, Overlapping use	Multi-device coordination patterns	5
Contextualized	Features x Subjective mediator/moderator	Subjective meaning of behavior	3, 4

Analytic Strategy

All empirical chapters employ multilevel modeling to account for the nested structure of experience sampling data, where repeated observations (level 1) are nested within individuals (level 2). This enables decompositions of variance into within-person and between-person components.

Prior to formal analysis, variables underwent several steps to allow for multilevel modeling. To isolate within-person from between-person differences, level 1 variables were person-mean centered (Enders & Tofighi, 2007). This ensures that fluctuations in an individual's behavior relate to their own mood or well-being, independent of stable between-person differences. Behavioral features were also within-person standardized (mean = 0, SD = 1, for each participant) to facilitate comparison across features with different scales.

Level 2 variables, representing person-level averages, were grand-mean centered, and primarily served to saturate the between-person model, rather than serve as a predictor (e.g., Chapter 3′s models).

To test temporally lagged associations (e.g., whether digital behavior at T0 predicts well-being at T1), selected variables were lagged to subsequent rows in the dataset. Constraints were applied to ensure valid temporal lagging: 1) within-person only, 2) within-day only, and 3) only between subsequent completed questionnaires. As such, we avoid lagging across different people, days, or questionnaires which were further back in time than the previous one.

While the dissertation focuses mostly on momentary associations, some analyses also examined daily associations as a comparison or robustness check. Momentary models used raw outcome variables (e.g., feeling happy, feeling rushed) measured at each ESM prompt, paired with the behavioral features aggregated within the corresponding time window. Daily models used aggregated daily averages for outcomes, and daily sums for predictors, which again were within-person centered and scaled (day totals within person totals). This approach lets us test whether associations differ depending on temporal granularity, an important topic in current media effects debates (Klingelhoefer et al., 2025).

Data processing and feature calculation were carried out in Python using pandas (team, 2020) for Chapters 3 and 4, Dask (Dask Development Team, 2016) for Chapter 3, and Polars (The Polars Development Team, 2025) for Chapter 5. Additional data transformations were carried out in R using dplyr (Wickham et al., 2025), esmpack (Viechtbauer & Constantin, 2023) and misty (Yanagida, 2023). Statistical analyses were also carried out in R, using: lme4 (Bates et al., 2015) in Chapters 4 and 5, lavaan (Rosseel, 2012) in Chapter 3, and in Python using pingouin (Vallat, 2018) in Chapter 4. All analysis code is available on the OSF page of each respective study.

Finally, each chapter's Methods section provides additional details on specific implementations concerning exact model specifications, handling of control variables, specific mediation or moderation tests, and model adaptations when fit criteria were not met.

Conclusion

This chapter has outlined the methodological foundations underlying the empirical studies presented in this dissertation. By combining intensive experience sampling with passively collected digital trace data across devices, the research design aimed to capture digital media use and digital well-being as dynamic, situated processes unfolding in everyday life.

Documenting the full processing pipeline makes clear how much decisions cascade. Choices made at early stages of the pipeline constrain and shape what is possible at later stages, in ways that are not always immediately visible. Consider the following chain: how we parse Screen data directly affects how we annotate Application data with screen sessions (Step 3 of application processing), which in turn helps determine the effective stop times and durations we calculate (Step 4). These durations become the basis for the volume and fragmentation features used later on. A different parsing decision at the screen level would have propagated through the entire pipeline, producing different duration and feature values, and hence potentially different inferential outcomes.

Similarly, notification cleaning decisions affect both the notification features used in Chapter 3 (responsiveness, reaction time) and the monitoring feature, which relies on screen activations without a preceding notification click. If our cleaning was overly strict, i.e. considering legitimate notifications as system noise, the monitoring metric would be systematically inflated, as more screen activations would appear to lack a linked notification.

These cascading dependencies are rarely acknowledged in digital trace research, where processing steps tend to be described more as independent operations. In practice, the pipeline functions in an interconnected manner, where early-stage decisions create dependencies that shape subsequent indicators. Documenting the full pipeline, as we have attempted here, is a necessary first step towards making these dependencies visible, but even our documentation cannot fully quantify their cumulative impact on the final analytic results.

The data processing pipeline documented in this chapter reveals that cleaning digital trace data is not a neutral, technical operation that separates signal from noise. Every cleaning decision entails a judgement about what constitutes relevant behavior. When we removed Always On Display or fingerprint sensor apps, or apps without package names, we were not correcting errors in an otherwise accurate record. We were deciding that these events should not be counted as screen time. When notification events were dropped from system packages, we judged these to be irrelevant to our notification-based indicators. The 240-minute outlier threshold for application durations entailed the assumption that single app events exceeding four hours more likely reflect a logging artefact than genuine behavior. Each of these decisions is defensible, but a different research team investigating other research questions could have easily reached different conclusions about what to retain or which steps to take.

Digital trace data are frequently described as providing a more accurate or objective account of behavior compared to self-reports (Parry et al., 2021). Our pipeline illustrates both the merit and the limits of this characterization. Trace data are indeed free from recall bias and social desirability effects. However, the events that a device logs are not the same as the behaviors a researcher wants to study. The gap between logged events and the theoretical constructs, what has been called the “inferential leap” (Conrad et al., 2021; as cited in Pankowska et al., 2025), must be bridged by a series of processing decisions, each of which introduces researcher judgement into the data. As such, these data occupy a position between objective and subjective. Recognizing this status does not diminish the value of trace data, but it also underlines the need to account for inherent biases and error present within these data (Bosch et al., 2025; Pankowska et al., 2025).

This chapter has aimed to be transparent about the trace-to-indicator pipeline. Transparency alone, however, is not the same as reflexivity. Throughout the preceding sections we have attempted to move between the former and the latter, by noting at various points how a different decision would have altered the resulting data. Yet, we recognize that our attempt remains incomplete. For every decision we have flagged and reflected on, others went unexamined, including decisions made by the tools themselves (e.g., how the logging app determines when an app “starts”). Full reflexivity would require much more space than we have already attributed. What we hope to have demonstrated is that the level of detail provided here already reveals a number of consequential choices that would otherwise remain invisible.

Several limitations of the current framework warrant discussion. First, logged data do not always allow for a clear distinction between non-use and missing data, especially when device-level restrictions (e.g., Android battery optimization or permission changes) interrupt logging without the researcher’s knowledge. Although we implemented monitoring and exclusion procedures to mitigate this, residual data loss likely remains in the dataset. Second, interaction with a device cannot be equated with engagement or attention. This is especially the case for desktop-based activities, where a window left open in the foreground may reflect passive presence rather than active use, a problem aggravated by the absence of AFK data. Relatedly, trace data cannot identify *who* is operating a device during a logged session. Device sharing has been recognized as a behavioral barrier to valid passive smartphone measurement more broadly (Keusch et al., 2022), though large-scale survey evidence suggests it occurs rarely in general adult populations. In our study, however, several participants reported lending their smartphone to their children, for instance to watch a video, meaning that these episodes were logged and attributed to the participant despite reflecting another person’s activity. This form of misattribution is largely invisible in the data and difficult to correct for systematically, though it could be partially addressed in future work through momentary self-report prompts or device-sharing flags. Third, the pipeline described here was developed for a specific dataset and research context. While we have attempted to document decisions in sufficient detail for others to evaluate and adapt, direct generalizability to other tools or research contexts should not be assumed.

These limitations highlight why behavioral trace data cannot replace subjective reports. Research questions related to how digital interactions are interpreted, whether experienced as stressful or meaningful, or how they fit within broader role demands, remain inaccessible through behavior logs alone. Experience sampling can hence greatly complement trace data and play a part in interpreting patterns of behavior more clearly.

The empirical chapters that follow build on this framework in complementary ways, and in doing so, test different aspects of the pipeline's assumptions and limitations. Chapter 3 examines the mismatch between subjective and objective indicators. Chapter 4 explores how increasing behavioral granularity and context alter associations with time pressure. Chapter 5 questions the consequences of extending measurement scope beyond the smartphone. Together, these studies illustrate how methodological decisions shape what digital trace data can and cannot tell us about digital media use and well-being.

Research Chapters

Connected Yet Cognitively Drained?

A Mixed-Methods Study Examining Whether Online Vigilance and Availability Pressure Promote Mental Fatigue[7]

[7]Published as: Van Gaeveren, K., Murphy, S. L., de Segovia Vicente, D., & Vanden Abeele, M. M. P. (2024). Connected Yet Cognitively Drained? A Mixed-Methods Study Examining Whether Online Vigilance and Availability Pressure Promote Mental Fatigue. *Communication Research*, 0(0). https://doi.org/10.1177/00936502241248494

Abstract

This mixed-methods study investigates whether online vigilance promotes mental fatigue, and whether this effect is greater when under pressure to be available online. Additionally, it examines whether passively sensed smartphone behavior can serve as a digital proxy for online vigilance. Data were collected from 1,315 adult participants, who received 84 experience sampling questionnaires over 14 days, providing 67,762 usable datapoints on individuals' perceptions of momentary online vigilance, mental fatigue, and availability pressure. Additionally, the smartphone use of 834 participants was passively monitored. Findings revealed both a momentary and lagged association between self-reported online vigilance and self-reported mental fatigue. Availability pressure was not a significant moderator, but did predict mental fatigue directly and indirectly, by promoting online vigilance. We found behavioral smartphone use features were weakly associated with self-reported online vigilance and mental fatigue. Overall, this study provides initial support that online vigilance may play a role in the development of mental health conditions such as burnout via its tendency to promote one of its precursors, mental fatigue.

Keywords: online vigilance, mental fatigue, availability pressure, smartphone, log data, passive sensing, monitoring, experience sampling

Introduction

Mental fatigue is a subjective state of tiredness characterized by a reluctance for further effort (Hopstaken et al., 2015). Attention towards mental fatigue has increased in recent years as research identifies the state as a key precursor to burnout (Demerouti et al., 2002), a mental health condition that has risen sharply in workers in westernized societies (De Hert, 2020; Edú-Valsania et al., 2022; Naldi et al., 2021). Together with other characteristics, such as reduced cognitive and emotional regulation and mental distancing (Schaufeli et al., 2020), mental fatigue is believed to contribute to feelings of exhaustion and a-motivation (Demerouti et al., 2002; Edú-Valsania et al., 2022). Given the debilitating health consequences associated with it, better understanding of what factors promote mental fatigue is essential to facilitate the development of effective interventions and policies.

Vigilance is a factor known to promote mental fatigue, as protracted periods of vigilance require hard mental work (Warm et al., 2008). For instance, truck driving (Ting et al., 2008), excavating (Li et al., 2019), and monitoring various security video feeds (Hodgetts et al., 2017) are mentally fatiguing activities because they demand paying constant attention to one or more stimuli. This study questions whether in contemporary Westernized societies a new type of vigilance may have emerged alongside the culture of ubiquitous digital connectivity, namely *online* vigilance, that may potentially contribute to rising burnout rates by promoting mental fatigue.

Online vigilance, then, refers to a state in which one is constantly attentive to the online world (Reinecke et al., 2018). Characterized by cognitive pre-occupation and a constant monitoring of and reactivity to the online world, we posit that online vigilance is a mentally taxing state; after all, performing day-to-day activities while simultaneously keeping constant tabs on what happens online, repeatedly checking various digital streams for incoming messages, responding immediately to incoming notifications, and making oneself available online at all hours of the day, may come with cognitive costs, including the costs of increased cognitive load (e.g., Gilbert et al., 2023) and of multitasking (e.g., switch costs, attention residue) (Freytag et al., 2021). Furthermore, because social pressures to be available online threaten autonomy (Halfmann & Rieger, 2019), and autonomy threats are fatiguing (Ryan & Deci, 2008), the impact of online vigilance on mental fatigue may be especially pronounced when perceiving greater pressure to be available online.

Drawing from a large-scale mixed-methods study (N = 1,315) that combines passive smartphone monitoring and experience sampling (67,762 data points), the aim of the present study is to test the idea that online vigilance promotes mental fatigue, especially when perceiving greater availability pressure. We test this hypothesis using both self-report and passively sensed behavioral measures of online vigilance, the latter helping to overcome validity and reliability concerns with self-report measures (e.g. Ellis et al., 2019; Parry et al., 2021), and potentially opening the door to the development of just-in-time-adaptive-interventions for early-onset detection of burnout (e.g., Wang & Miller, 2020).

Theoretical Framework

Online Vigilance as the Subjective Experience of Being 'On'

Ubiquitous connectivity serves obvious purpose and utility, making communication possible at great distances, enabling immediate satisfaction of various needs and desires, and offering various novel functionalities. But with these opportunities an 'always on' culture has emerged (McDowall & Kinman, 2017; M. M. P. Vanden Abeele, 2021), in which people are 'permanently online, permanently connected' (Vorderer et al., 2018), and find it increasingly difficult to switch off (Nguyen, 2021, 2023). Recent research shows that a repercussion of this 'always on' lifestyle is the tendency for individuals to be highly attuned to their online world, i.e., to be online vigilant (Johannes, Meier, et al., 2021; Reinecke et al., 2018).

Online vigilance is argued to comprise three distinct components (Reinecke et al., 2018): *salience*, which reflects a cognitive preoccupation with the online world (e.g., thinking about incoming emails, Twitter feed, WhatsApp messages), *reactibility*, which reflects how quickly one responds to and prioritizes events and cues from the online sphere relative to offline demands, and *monitoring*, which reflects how often one actively enters their online sphere to check for new updates and information. Although online vigilance can be understood as a trait characterizing between-person differences in people's mindsets towards the online world (Reinecke et al., 2018), ample research also supports its state-like nature: Studies demonstrate that people vary substantially moment-to-moment in online vigilance, and this variability correlates with various well-being outcomes (e.g., Johannes et al., 2019; Freytag et al., 2021; Gilbert et al., 2023; Johannes, Meier, et al., 2021; Reinecke et al., 2018). The central hypothesis guiding this study is that one such well-being outcome affected by online vigilance is mental fatigue.

Online Vigilance as a Precursor to Mental Fatigue

The claim that online vigilance elevates individuals' experiences of mental fatigue is grounded in a strong theory and evidence base in the field of psychology on the cognitive cost of being in a state of vigilance. Vigilance, a term often used interchangeably with sustained attention, can be understood as "the ability of observers to maintain their focus of attention and remain alert to stimuli over prolonged periods of time" (Warm et al., 2008, p. 115). Although originally believed to require minimal mental effort (Grier et al., 2003; Warm et al., 2008), recent empirical literature concludes that vigilance promotes fatigue — and does so frequently and sometimes to a non-minor extent (Finomore et al., 2006; Matthews et al., 2010; Smith et al., 2019; Warm et al., 2008), by depleting cognitive resources (Eisert et al., 2016). In line with resource expenditure theory (i.e. overload theory) prolonged vigilance is found to increase cognitive load, thereby leading to a faster and greater depletion of cognitive resources (Head & Helton, 2014). This depletion becomes even more pronounced when multitasking is involved, as dual-task performance requires allocating additional cognitive resources to the control of attention over two (or more) stimuli (Courage et al., 2015).

Applying this evidence and reasoning to the online sphere, it seems plausible that constantly attuning oneself to the online sphere, i.e., being *online* vigilant, is fatiguing. A key ingredient here is the 'always on' yet fragmented nature of online communication (Oulasvirta et al., 2012; Reinecke et al., 2018). To remain fully abreast of content and notifications — which often update moment-to-moment and become outdated just as quickly — one needs to keep the online sphere top of mind, as well as monitor and potentially immediately react to incoming information or communication. The latter activities require alternating attention from the ongoing task/activity to the digital sphere and back (Y. Peng & Tullis, 2021). As such, individuals may face a depletion of cognitive resources not just because of being attentive to what is happening online increases the overall cognitive load, but also from the capacity costs associated with switching attention to and from communication channels, a cognitively depleting experience that has been well-documented in the literature on media multitasking (e.g., Baumgartner & Sumter, 2017; Rioja et al., 2023).

Circumstantial support for the fatiguing effect of online vigilance is visible, among others, in research on the work of journalists, who identify the constant monitoring of social media spaces as mentally taxing (Bossio & Holton, 2021). There is also a growing body of evidence on online vigilance that indirectly supports its fatiguing nature: Freytag et al. (2021) found that cognitive preoccupation with online communication (i.e., the salience component of online vigilance), together with media multitasking, depletes working memory and situational coping capacities. Similarly, Gilbert et al. (2023) found that online vigilance increases people's communication load, and Johannes, Meier, et al. (2021) found that online vigilance predicted a measure of affective wellbeing that included feeling tired (vs. awake) as one of its indicators. To date, however, research linking online vigilance directly to the experience of mental fatigue is still lacking. Moreover, to our knowledge, studies have not yet explored whether the effect of online vigilance lingers, which would strengthen causal claims over online vigilance and its potential to accumulate over time, thus perhaps contributing to the development of feelings of severe exhaustion, as witnessed in the case of burnout. Extant findings on the effect of online vigilance also result from studies involving student-heavy samples (e.g., Freytag et al., 2021; Gilbert et al., 2023). Accordingly, the first objective of this study is to bridge these knowledge gaps by formally testing the contemporaneous and lagged associations between online vigilance and mental fatigue in an adult sample:

> H1: Online vigilance positively predicts mental fatigue at the same (H1a) and the next time-point (H1b).

Availability Pressure as Moderator

Digital communication comes with normative expectations attached (Bayer et al., 2016), among others expectations to be 'always on' (Nguyen, 2021). Individuals can perceive these expectations in the form of social pressure to be available to others online (Halfmann & Rieger, 2019). Prior research suggests that such perceived pressure exacerbates the negative effect of online vigilance on well-being outcomes. Gilbert et al. (2023), for instance, found that the negative impact of online vigilance on stress is greater when availability pressure is high. They explain this effect through the lens of self-determination theory, a theory that gives a comprehensive account of how internal and external pressures can rid individuals of their sense of autonomy (Ryan & Deci, 2008). Halfmann et al. (2021; see also Halfmann & Rieger, 2019) already observed that availability pressure can frustrate individuals in their autonomy needs. Building further on this observation, Gilbert et al. (2023) argued that under conditions of heightened availability pressure, people likely appraise the stress response resulting from online vigilance differently as they experience reduced autonomy to cope with it; Therefore, the debilitating effect on stress may be greater.

A similar pattern may exist for mental fatigue. First of all, autonomy-frustrating experiences are themselves often judged as fatiguing (e.g., Sheldon et al., 1996). Hence, if we consider availability pressure an autonomy-frustrating boundary condition, availability pressure itself may elicit fatigue. Organizational literature provides evidence for this, showing perceived expectations to be available after-hours to predict burnout symptoms including exhaustion (Hendrikx et al., 2023) and a more negative work–life balance (Belkin et al., 2020).

Because availability pressure in itself leads to a depletion of resources, however, situations of heightened pressure may also lead to an overall diminished capacity to regulate behavior so that energy is maximally preserved; circumstantial evidence to support this reasoning was found by J. Hall (2017). He observed that 'mobile entrapment', a construct similar to availability pressure, negatively impacts a measure of subjective well-being comprising items reflecting how energetic and fatigued individuals feel. Moreover, mobile entrapment also may act as a boundary condition that weakens the positive effect of texting on well-being, leading Hall to conclude that availability pressure shapes individual experiences. In the context of online vigilance and mental fatigue, the resource depletion resulting from heightened availability pressure may mean that individuals have less resources left to cope with and protect themselves against the fatiguing effect of online vigilance. The second objective of this study is to test this idea, by examining whether the association between perceived online vigilance and subjective fatigue is moderated by perceived availability pressure (H2a), and whether this relation holds true when looking at lagged effects (H2b).

> H2: Availability pressure exacerbates both the momentary (H2a) and lagged (H2b) positive association between online vigilance and mental fatigue.

Behavioral Measures and Objective Assessments of Online Vigilance

To date, studies examining online vigilance have predominantly relied on self-report data. However, online vigilance can arguably be measured behaviorally, by examining how quickly people respond to notifications, or how often they check or pick up their smartphone, as these behavioral tendencies seem likely to capture this sense of vigilance to one's online world (Johannes, Meier, et al., 2021; Reinecke et al., 2018).

There is clear value in the development of behavioral online vigilance indicators, as they open up the possibility of implementing just-in-time interventions to regulate excess vigilance through passive sensing. Also, they overcome limitations of self-report measures of online vigilance, as we may assume it is difficult for people to accurately recall or capture how vigilant they were over the past few hours.

To our knowledge, only Johannes, Meier, et al. (2021) has explored whether objective assessments of smartphone use behavior match self-perceptions of the behavioral components of online vigilance. Unfortunately, due to technical issues the authors used secondary operationalizations for which a theoretical link was less intuitive, and their study therefore did not show convincing evidence. This leaves room for new analyses that explore the potential behavioral correlates of online vigilance. Hence, a final objective of this study is to explore whether two of the behavioral components of perceived online vigilance, reactibility and monitoring, have behavioral correlates in smartphone use (RQ1), and whether these correlates perform similarly in predicting mental fatigue.

> RQ1: Can passively sensed behavioral smartphone use patterns serve as behavioral proxies for online vigilance, more specifically for reactibility and monitoring?
>
> RQ2: How do these behavioral patterns perform in predicting momentary and temporally lagged fatigue?

In line with these objectives, we will extend our investigation by assessing our four hypotheses H1a, H1b, H2a, and H2b, using behavioral measures of online vigilance in addition to self-report measures. These hypotheses represent H3a, H3b, H4a, and H4b, respectively.

Method

Participants

This preregistered study uses data gathered from 1,315 adult participants (mean age = 38.83 years; *SD* = 11.72), who participated voluntarily in a two-week mixed-methods study in the Fall of 2022. The study took place in the context of a citizen-science project, and was advertised with the aid of a large newspaper (with a daily reach of around 450,000 readers) who invited people to participate in our study through both their print and online news channels (see @#sec-appendix-a for further details). The project received approval from the institutional review board of Ghent University.

The 1,315 participants included in formal analyses provided 67,762 ESM datapoints, which equates to an average of 51.53 questionnaires completed per participant. Most participants (62%) identified as female (37% male; 1% non-binary/other), were highly educated (93% had at least a bachelor's degree), and were actively working (77%; 8% were students; 15% were not in paid employment, e.g., retired, sick leave, etc.). Most participants used an Android device (70%), a percentage higher than the general adult population in this region (62% in 2021 (Sevenhant et al., 2022)). The remaining 30% used an iOS device.

Although 916 participants were Android users and installed the tracking app so their smartphone could be logged, due to various technical difficulties (e.g., smartphone operating system blocking tracking access, see @#sec-appendix-a for more information), log data could only be collected from 834 participants (91%). More than 7 million unique phone-events were collected from these participants, comprising their app, screen, notification and keyboard activity.

Procedure

All participants were asked to install m-Path, an app to receive experience sampling questionnaires. Android users were additionally invited to install a dedicated app built for the project, that passively monitored participant smartphone behavior[8] (see below for further details). See the section 'Behavioral measures' for details.

In the experience sampling app, over a period of 14 days participants received 6 questionnaires per day between 7.30 am and 10.45 pm. Each notification was randomly scheduled within a 90 minute timeslot, and remained available for 45 minutes (with a reminder after 30 minutes). Time intervals varied from minimally 1 hour and 15 minutes to maximally 4 hours and 15 minutes between two subsequent questionnaires; on average responses were separated by 3 hours and 4 minutes (*SD* = 16.66 minutes). Additional details about the onboarding procedure and the ESM questionnaire can be found in @#sec-appendix-a.

Measures

3.4.c.a Self-report Measures

The ESM items relevant to the present study were measured on a 7-point Likert scale ranging from 1 ('Not at all') to 7 ('All the time'). The first questionnaire of the day had items that referred to the time period since waking up (e.g., "Since getting up this morning, I felt […]"), while all other questionnaires had items that referred to the time period since the previous questionnaire they received (e.g., "Since the last questionnaire, I felt […]").

Online Vigilance. The online vigilance measure included three items that were adapted from Reinecke et al. (2018), reflecting monitoring ('[…] I continuously monitored my emails and messages.'), reactibility ('[…] I gave incoming messages and emails my immediate attention.'), and salience ('[…] my thoughts were continuously with my emails, messages and social media.').

Availability Pressure. Availability pressure was measured with one item, stating '[…] I felt pressure to be digitally available to others.'

[8] This app functions solely on Android devices and, to our knowledge, no iOS alternatives exist which can capture detailed trace data.

Mental fatigue. Mental fatigue was measured with one item, stating '[…] I felt mentally drained.'

Wake-up time. To mark the onset of the first time window of the day, wake-up time was also measured in the first questionnaire of each day with the following item (time-entry field): 'At what time did you get up today (hh:mm, e.g., 07:30)?'

3.4.c.b Behavioral Measures

We passively monitored Android users' smartphone use, gathering information on start and stop times of smartphone sessions and app activities, as well as timestamped notification data. Three digital proxies were calculated from these data, with one of these features representing the monitoring dimension of online vigilance, and two of them representing the reactibility dimension of online vigilance. These features were computed for each ESM time interval. For instance, if the fourth ESM questionnaire of the day asked participants how mentally fatigued they felt since the previous questionnaire they received, each digital proxy of online vigilance covered that exact same period to facilitate formal analyses. The time window for each survey corresponded to the period from the sending of the previous survey, as clarified in both the participation instructions and the baseline survey.

Monitoring Behavior. To capture the monitoring dimension, we computed a feature reflecting the number of times a participant unlocked and relocked their phone again within 5 seconds, without directly reacting to/accessing the phone via a notification pop-up. This action can be seen as a "self-initiated" or voluntary phone unlock that merely serves the purpose of 'checking'.

Responsiveness. To capture the reactibility dimension, we computed a first feature representing the percentage of notifications reacted upon, calculated as the proportion of (a) number of notifications clicked (= reacted upon) divided by (b) number of notifications received.

Reaction Time. Also, to capture the reactibility dimension, we computed a second feature reflecting the time it took to react to notifications, calculated by computing the average amount of time between receiving an app notification and clicking it.

These last two features were computed based on the notification data from communication and social media apps (e.g., e-mail, (work) communication, social media apps). Further information on data cleaning and processing can be found in @#sec-appendix-a.

Data Analysis

ESM data is nested (observations within participants). Thus, a multilevel model was required to prevent standard error underestimation (Hox, 1998). Also, given that our models include latent factors, SEM was required to model associated errors. To meet both requirements, a multilevel SEM was conducted. To enable these analyses using the R package lavaan, level 2 of our model was saturated (see https://lavaan.ugent.be/tutorial/multilevel.html for details). This step also ensured model fit indices concerned level 1 of our model, our primary focus. As part of these analyses, random intercepts were used with fixed slopes. Prior to testing our structural model, however, we conducted confirmatory factor analysis (CFA) to test our measurement model for online vigilance, both as indicated by self-report and by behavioral variables. Model fit indices were used to assess our structural models given present degrees of freedom, with adequate fit determined based on the following three criteria: (a) *CFI*: above .90: moderate but acceptable model fit, above .95: very good model fit. (b) *RMSEA*: lower than .06 good model fit, between .06 and .08 moderate but acceptable model fit, and above 0.08 poor model fit. (c) *SRMR*: .05 or below demonstrates good model fit. These criteria align with current standards (Hu & Bentler, 1999; MacCallum et al., 1996; Marsh et al., 2004). If model fit criteria could not be met, our models were adapted in line with theory and output modification indices. Lastly, we employed two-tailed testing, and considered effects significant if p-values were .05 or less.

Transparency and Openness

Information about our preregistration, code and analysis, as well as the larger project in which this study is situated, can be found on osf (https://osf.io/4hb2c/). Our preregistration stated that when structural models would not reach adequate fit, theoretically and methodologically relevant adaptations would be implemented. Unfortunately, this was the case for all but one of our models. We specify at the start of the confirmatory results section which adaptations were made, and for which reasons.

Results

Descriptives

Descriptives for all measures per ESM time window and for daily smartphone use are included in Table 3.1. On average participants were very rarely to rarely vigilant to their online world (self-report variables), under pressure to be available online, or experiencing mental fatigue. Participants spent on average around 180 minutes on their smartphone per day, opened approximately 107 apps, unlocked their devices 85 times and received about 41 notifications each day.

Table 3.1: Summary of the person-level mean values for each uncentered or unstandardized variable, and for daily smartphone use.

	n	Mean	*SD*	Min	Max
Person-level means descriptives[a]					
salience[s]	1315	2.45	0.94	1.00	6.24
monitoring[s]	1315	2.75	0.98	1.00	6.16
reactibility[s]	1315	3.08	1.01	1.00	6.51
mental fatigue[s]	1315	2.83	1.20	1.00	6.86
availability pressure[s]	1315	2.45	1.04	1.00	6.62
monitoring sessions[l]	767	9.61	6.36	1.00	59.00
responsiveness[l] (%)	663	0.30	0.18	0.01	1.00[b]
reaction time[l] (sec.)	655	641.20	455.58	0.71	3565.61
Daily smartphone use descriptives[c]					
screentime (min.)	702	179.80	80.98	11.59	502.18
app frequency	702	106.90	54.52	8.60	337.08
screen unlocks	761	84.77	42.00	3.00	262.14
notifications	736	33.62	29.32	1.57	261.33

Note. s denotes self-report variables; l denotes behavioral variables. a Person-level means per ESM time window. b One participant's average was excluded due to a technical issue with their data. c Daily smartphone descriptives were measured for participants with at least seven days of data.

Self-report Models

We first report the results of the self-report models. The analyses for these models were based on data from the total sample of 1,315 adults. As mentioned above, we deviated from our preregistered analysis plans when model fit was unsatisfactory. We did so in three ways:

First, we included covariances between variables measured at the same time-point as model fits improved greatly when doing so, and this also makes theoretical sense. Second, for the temporally lagged models, we added a cross-lagged path from mental fatigue (T0) to online vigilance (T1). We did so because model fit of the original model was poor, whereas the cross-lagged model fitted well and, in our view, only provides greater support for the hypotheses tested. Third, theory indicated a direct impact of availability pressure on mental fatigue, which we had not preregistered. Again, adding this effect greatly improved model fit. Hence, we adapted our models accordingly. To keep the results section concise, only the results of the finalized models are discussed below. The original models (with poor model fit) are added in @#sec-appendix-a, and the step-by-step procedure is detailed on osf.

3.5.b.a *Confirmatory Factor Analysis of Self-Report Online Vigilance Items*

We first performed a confirmatory factor analysis to test whether our three online vigilance items loaded onto one factor[9]. The (just-identified) model showed that the factor loadings of the salience, monitoring and reactibility self-report items were significant with β's ranging from .68 to .81 (see Table 3.2). Additionally, these items demonstrated high within-person (r = .50 to .59) and between-person (r = .82 to .89) correlations. These findings (also detailed in Table 3.3) broadly support using a one-factor structure for the self-report measure of online vigilance.

Table 3.2: Factor loadings.

Latent Factor	Indicator	*B*	*SE*	*Z*	*p*	*β*
Self-reported Online Vigilance						
	salience	1.00	.00			.68
	monitoring	1.36	.01	148.65	<.001	.81
	reactibility	1.26	.01	149.75	<.001	.73
Behavioral Online Vigilance						
	responsiveness	1.00	.00			.23
	monitoring sessions	−4.02	2.95	−1.36	.173	−.83
	reaction time	.35	.08	4.67	<.001	.08

Table 3.3: Within- and between-person correlation matrix.

	1	2	3	4	5	6	7	8
1. salience[s]	—	.89***	.82***	−.10***	−.16***	.14***	.43***	.84***
2. monitoring[s]	.55***	—	.90***	−.12***	−.10***	.09***	.34***	.79***
3. reactibility[s]	.50***	.59***	—	−.11***	−.14***	.12***	.30***	.74***
4. reaction time[l]	−.03***	−.05***	−.06***	—	.02***	−.05***	−.08***	−.07***
5. responsiveness[l]	−.07***	−.09**	−.09***	.03*	—	−.15***	−.09***	−.16***
6. monitoring sessions[l]	.09***	.10***	.10***	−.07***	−.19***	—	.05***	.12***
7. mental fatigue[s]	.17***	.17***	.14***	−.01	−.03*	.02***	—	.46***
8. availability pressure[s]	.45***	.47***	.44***	−.03***	−.11***	.09***	.18***	—

Note. * $p < .05$, ** $p < .01$, *** $p < .001$. Within-person correlations are shown below the diagonal; between-person correlations above the diagonal. These correlations do not take the

[9]Because this model was just-identified, model fit could not be examined using standard model fit indices (e.g., CFI, RMSEA, etc.; Hu & Bentler (1999); MacCallum et al. (1996); Marsh et al. (2004)).

multilevel structure into account, and hence will be slightly inflated. s denotes self-report variables; l denotes behavioral variables.

3.5.b.b *Hypotheses Tests with Self-reported Online Vigilance*

Hypothesis 1 posited that self-reported online vigilance would predict mental fatigue, both at the same timepoint (H1a) and lagged (H1b). Hypothesis 1a was supported by the data (see Table 3.4), showing a clear effect of online vigilance on mental fatigue ($\beta = 0.17$, $p < .001$), controlling for mental fatigue at the previous time-point. The model fit the data well. Hypothesis 1b was also supported (see Table 3.4), with online vigilance predicting greater mental fatigue at the next time-point ($\beta = .07$, $p < .001$), controlling for mental fatigue at the present time-point ($\beta = .39$, $p < .001$). It should be noted that we also found a cross-lagged effect of mental fatigue on online vigilance ($\beta = .02$, $p < .001$).

We performed an additional (non-preregistered) exploratory analysis to examine whether there was the effect of online vigilance on mental fatigue lagged even further (lag-2 effect). Controlling for lag-1 online vigilance, we found that lag-2 online vigilance also predicted mental fatigue ($\beta = .01$, $p = .049$), thus showing that the fatiguing effect of online vigilance lingers on, leaving a residue at least 4 to 6 hours later.

Hypothesis 2 stated that availability pressure would moderate the aforementioned contemporaneous (H2a) and temporally lagged (H2b) associations, so that in instances where availability pressure was higher, the association between online vigilance on mental fatigue would be stronger. The hypothesis was not supported (see Table 3.4): availability pressure did not significantly moderate the contemporaneous ($\beta = .01$, $p = .631$), nor the lagged ($\beta = .01$, $p = .666$) association between online vigilance and mental fatigue.

When fitting the moderation models, we observed that availability pressure had a direct effect on mental fatigue in both the contemporaneous and lagged models, and that it also covaried strongly with online vigilance. A similar observation was made in recent work from Gilbert et al. (2023), in which the authors suggest to explore mediation models. Based on our observations and following this suggestion, we performed an additional exploratory analysis, modelling a lagged effect of availability pressure on online vigilance, calculating both its direct lagged and its indirect effect (via online vigilance) on mental fatigue. Findings support a direct lagged effect of availability pressure on mental fatigue ($\beta = 0.06$, $p < .001$), and an indirect effect via online vigilance ($\beta = 0.05$, $p < .001$; see Table 3.5).

Table 3.4: Regression coefficients, covariances and model fit — Self-report models.

Hypothesis	Regressions	Indicator	*B*	*SE*	95% CI	*p*	*β*	CFI	RMSEA	SRMR
H1a — Outcome: mental fatigue										
	Predictor	online vigilance	.25	.01	[.24, .27]	<.001	.17	.99	.05	.05
		lagged mental fatigue	.38	.00	[.38, .39]	<.001	.39			
H1b — Outcome: mental fatigue										
	Predictor	lagged online vigilance	.10	.01	[.08, .11]	<.001	.07	.96	.09	.08
		lagged mental fatigue	.39	.00	[.38, .40]	<.001	.39			
	Predictor (online vigilance)	lagged online vigilance	.48	.01	[.47, .50]	<.001	.49			
		lagged mental fatigue	.02	.00	[.01, .02]	<.001	.02			
	Covariance	mental fatigue — online vigilance	.12	.00	[.12, .13]	<.001	.17			
		lagged mental fatigue — lagged online vigilance	.20	.01	[.19, .21]	<.001	.21			
H2a — Outcome: mental fatigue										
	Predictor	online vigilance (ov)	.17	.01	[.15, .19]	<.001	.13	.99	.04	.05
		lagged mental fatigue	.38	.00	[.37, .39]	<.001	.39			
		availability pressure (ap)	.05	.01	[.03, .08]	<.001	.05			
		ov × ap	.00	.00	[−.01, .01]	.631	.01			
	Covariance	online vigilance — availability pressure	.58	.01	[.56, .59]	<.001	.60			
H2b — Outcome: mental fatigue										
	Predictor	lagged online vigilance (ov)	.06	.01	[.04, .09]	<.001	.04	.99	.04	.06
		lagged mental fatigue	.39	.00	[.38, .39]	<.001	.39			
		lagged availability pressure (ap)	.03	.01	[.01, .05]	.008	.03			
		ov_lagged × ap_lagged	.00	.00	[−.01, .01]	.666	.01			
	Predictor (online vigilance)	lagged online vigilance	.49	.01	[.47, .50]	<.001	.49			
		lagged mental fatigue	.02	.00	[.01, .03]	<.001	.03			
	Covariance	online vigilance — mental fatigue	.12	.00	[.11, .13]	<.001	.17			
		lagged online vigilance — lagged mental fatigue	.05	.00	[.05, .06]	<.001	.06			

Table 3.5: Mediation model.

Regressions	Indicator	*B*	*SE*	95% CI	*p*	*β*	CFI	RMSEA	SRMR
Outcome: mental fatigue									
	availability pressure	.06	.01	[.05, .08]	<.001	.06	.99	.03	.01
	lagged availability pressure (direct)	.06	.01	[.05, .07]	<.001	.06			
	lagged availability pressure (indirect)	.06	.00	[.05, .06]	<.001	.05			
	online vigilance	.24	.01	[.22, .26]	<.001	.16			
Outcome: online vigilance									
	lagged availability pressure	.23	.00	[.23, .24]	<.001	.33			
Covariance									
	availability pressure — online vigilance	.45	.01	[.44, .47]	<.001	.54			
	availability pressure — lagged availability pressure	.40	.01	[.39, .42]	<.001	.32			

3.5.b.c *Exploratory Day-Level Analysis*

Finally, we added an exploratory analysis to examine whether the associations between availability pressure and online vigilance uphold at the day-level. This was the case (see Table 3.6): Average daily availability pressure ($\beta = 0.15, p < .001$) and online vigilance ($\beta = 0.17, p < .001$) predicted average daily mentally fatigue; there was an indirect effect of availability pressure via online vigilance ($\beta = 0.12$, $p < .001$).

Table 3.6: Exploratory day-level model.

Regressions	Indicator	*B*	*SE*	95% CI	*p*	*β*	CFI	RMSEA	SRMR
Outcome: daily mental fatigue									
	daily availability pressure	.17	.01	[.14, .19]	<.001	.15	.99	.05	.01
	daily availability pressure (indirect)	.14	.01	[.12, .15]	<.001	.12			
	daily availability pressure (total)	.30	.01	[.29, .32]	<.001	.27			
	daily online vigilance	.25	.02	[.21, .28]	<.001	.17			
Outcome: daily online vigilance									
	daily availability pressure	.55	.01	[.54, .57]	<.001	.70			

Behavioral Models

3.5.c.a Confirmatory Factor Analysis of Behavioral Online Vigilance Items

The next step in our analysis was to explore whether we could infer behavioral proxies for online vigilance from the smartphone log data. We first performed a CFA with the three behavioral indicators that we calculated, reflecting monitoring (operationalized as pro-active checking of the phone) and reactibility (operationalized as relative responsiveness to notifications and response speed). Interestingly, factor loadings did not support the features loading onto one factor, with β's ranging from .08 to −.83 (see Table 3.2), and unexpectedly low within- ($r = .03$ to $-.19$) and between-person ($r = .02$ to $-.15$) correlations (see Table 3.3). We further explored associations between the self-report online vigilance items and the behavioral features, and here also, correlations were surprisingly weak at both the within (r = .04 to .07) and between-person level (r = .09 to −.16; see also Table 3.3).

Although, based on the evidence, we cannot claim that the selected features serve as proxies for the experience of online vigilance, we nonetheless examined each feature's capacity to predict mental fatigue, both contemporaneously and lagged, and with and without availability pressure as a moderator.

3.5.c.b Hypotheses Tests with Behavioral Online Vigilance Measures

Parallel to our first hypothesis, our third hypothesis posited that, controlling for previous fatigue levels, behavioral online vigilance would predict mental fatigue, both contemporaneously (H3a) and lagged (H3b). We tested models for each separate feature. These base models were just-identified, hence their fit could not be examined. After modelling covariances as free parameters, however, both models showed good fit (see Table 3.7). No support for an effect was found.

Adding availability pressure as a moderator in the contemporaneous (H4a) and temporally lagged (H4b) models yielded similar results, with one exception for reaction time, which showed a small but significant, and counter-intuitive contemporaneous effect (β = .07, p = .048) on mental fatigue after controlling for availability pressure and the interaction effects. This finding suggests that when people respond more slowly to notifications, they are more fatigued. While theoretically not implausible, overall, these results lead us to reject all four behavioral hypotheses (H3a to H4b).

Table 3.7: Regression coefficients and model fit — Behavioral models.

Hypothesis	Regressions	Indicator	*B*	*SE*	95% CI	*p*	*β*	CFI	RMSEA	SRMR
H3a — Outcome: mental fatigue										
	Predictor	reaction time	.00	.02	[−.03, .04]	.835	.00	1.0	.00	.01
		responsiveness	−.03	.02	[−.07, .01]	.086	−.02			
		monitoring sessions	.01	.02	[−.03, .04]	.630	.01			
		lagged mental fatigue	.39	.01	[.36, .41]	.000	.39			
H3b — Outcome: mental fatigue										
	Predictor	lagged reaction time	−.01	.02	[−.05, .03]	.613	−.01	.99	.02	.02
		lagged responsiveness	−.01	.02	[−.05, .03]	.659	−.01			
		lagged monitoring sessions	.01	.02	[−.03, .05]	.541	.01			
		lagged mental fatigue	.38	.02	[.35, .41]	.000	.38			
H4a — Outcome: mental fatigue										
	Predictor	reaction time	.10	.05	[0, .19]	.048	.07	.99	.05	.05
		responsiveness	−.01	.04	[−.1, .07]	.738	−.01			
		monitoring sessions	.05	.04	[−.03, .12]	.215	.04			
		lagged mental fatigue	.37	.01	[.35, .4]	.000	.38			
		availability pressure (ap)	.19	.03	[.13, .25]	.000	.18			
		reaction time × ap	−.09	.05	[−.19, .01]	.067	−.07			
		responsiveness × ap	.00	.05	[−.09, .09]	.961	.00			
		monitoring sessions × ap	−.06	.04	[−.14, .03]	.206	−.05			
H4b — Outcome: mental fatigue										
	Predictor	lagged reaction time	.05	.05	[−.04, .13]	.316	.04	.98	.07	.06
		lagged responsiveness	−.02	.04	[−.1, .06]	.654	−.01			
		lagged monitoring sessions	.01	.04	[−.06, .09]	.715	.01			
		lagged mental fatigue	.37	.02	[.34, .4]	.000	.37			
		lagged availability pressure (ap)	.10	.03	[.04, .16]	.000	.09			
		lagged reaction time × ap	−.06	.05	[−.15, .03]	.199	−.05			
		lagged responsiveness × ap	.02	.04	[−.06, .11]	.587	.02			
		lagged monitoring sessions × ap	−.01	.05	[−.1, .08]	.831	−.01			

Discussion

Psychological research on *offline* contexts provides ample evidence that state vigilance promotes mental fatigue, yet to our knowledge, this study is among the first to investigate whether *online* vigilance affects individuals similarly. This is an important question to investigate. After all, online vigilance has already been found to predict stress (Freytag et al., 2021) and reduce affective well-being (Johannes, Meier, et al., 2021). If it also mentally fatigues individuals, then this further justifies societal concerns over the harms associated with being 'always on'. Most notably, given the strong link between mental fatigue and burnout (Edú-Valsania et al., 2022), knowing that online vigilance promotes mental fatigue brings immediate implications for how we understand and tackle this mental health condition.

Drawing from the data of 1,315 adult individuals, we investigated this question by exploring the momentary and lagged impact of *online* vigilance on mental fatigue, and additionally examined whether the pressure to be constantly available might exacerbate this link. Finally, we explored the potential of behavioral smartphone features to serve as 'passively sensed' proxies for online vigilance, which would open the door to detect the early onset of mental fatigue.

Overall, our self-report findings lend strong support for the hypothesis that online vigilance promotes mental fatigue, both momentarily and lagged: constantly attending to, monitoring and reacting to the online world comes with an energy-sapping cost, not only 'in the moment' but also further in time. In line with psychological theories on vigilance (e.g., Warm et al., 2008), we may thus conclude that online vigilance represents a form of 'mental work' with fatiguing qualities. This fatiguing effect may linger in time because fatigued individuals may lack the resources necessary to actively regulate themselves (see also Inzlicht et al., 2014). With mental fatigue also predicting online vigilance, there may thus be a risk for mutual reinforcement that could ultimately lead to a negative spiral of fatigue. Together with the extant evidence on other maladaptive outcomes of online vigilance such as stress (e.g., Freytag et al., 2021; Gilbert et al., 2023; Reinecke et al., 2018), our findings support broader concerns over being 'always on' and its repercussions for individual health and well-being.

Important to note here, is that the magnitude of the identified effects here is small; an individual would have to become *much* more vigilant of their online world (e.g., if an individual moves from stating they 'very rarely' were vigilant of their online world in the past hours, to stating they were 'very often' vigilant of it) to move even 1-point upwards on the mental fatigue scale (e.g., from 'rather true' to 'true'). Yet, although the effect is very small in magnitude, and intriguingly may be so small that individuals may not easily perceive it, the lagged associations do indicate that mental fatigue effects may not get fully regulated, and could instead accumulate over-time, similar to how small calorie increases can snowball into more consequential downstream effects (e.g.: Most et al., 2017). Hence, small effects per se should not be discounted out of hand, particularly when effects pertain to very narrow timescales (e.g., a few hours). That stated, further research is needed before concluding a pronounced impact on longer-term outcomes such as burnout.

A second aim of our study was to examine whether availability pressure exacerbates the impact of online vigilance on mental fatigue. After all, self-determination theory suggests internal or external pressures have a depleting quality (see also Ryan & Deci, 2008), hence it would be plausible that under greater availability pressure individuals lack the necessary resources to cope with and regulate the fatiguing effect of online vigilance effectively. No moderating effect was found, however, although availability pressure did predict mental fatigue directly (thus supporting its depleting effect), both acutely and downstream (i.e., next time-point) and also indirectly, through online vigilance. This result shows a remarkable parallel with the findings from Gilbert et al. (2023) recent study exploring the link between online vigilance, and we agree that “studies with an even more granular temporal level, or experimental studies are needed that are able to clearly delineate the various roles these concepts play” (p. 452). A tentative explanation for the nonsignificant moderating effect that could be explored in such future research, is the proposition that availability pressure itself might not always be negatively valenced; for instance, one might perceive that others expect them to be available in anticipation of good news (e.g., a birth announcement). Such availability expectations might not necessarily reduce or frustrate autonomy needs, but may on the contrary reflect belongingness, which may in turn drive individuals to remain vigilant over their communication channels.

A third and final aim of this study was to examine whether three behavioral smartphone use features which we believed would give expression to the behavioral dimensions of online vigilance, namely the speed at which an individual responds to notifications (reactibility), the percentage of notifications they responded to (reactibility), and the amount of self-initiated smartphone unlocks (monitoring), would predict mental fatigue. The value of these features lies in their possibility to be passively monitored, which overcomes issues associated with self-report data such as recall bias and common method variance, but also opens the door to digitally phenotyping online vigilance, and building automatized interventions on those digital phenotypes (e.g., sending the user a prompt, informing them they might benefit from a pause). Unexpectedly, however, our behavioral features appeared a poor measure of self-reported online vigilance: They showed a very weak association between themselves, as well as with the corresponding self-reported online vigilance items. Likely as a result of the latter, they did not perform well in predicting mental fatigue.

Various explanations can be sought for this unexpected finding. A first obvious possibility is that the chosen features have low validity: Perhaps they do not discriminate well between state online vigilance and, alternatively, the normal waxing-and-waning of task engagement. After all, human beings have an innate motivation to balance ‘exploitation and exploration’ (Inzlicht et al., 2014) — i.e., to alternate task engagement with the pursuing of alternative, potentially online activities (Wiradhany et al., 2021). For instance, short social media breaks may help individuals recover from a taxing work task. It may be that, at a behavioral level, the features that we identified as representations of online vigilance may share overlap with what smartphone behavior looks like when relaxing from work, creating a major confound in the association between these behaviors and the subjective experience of online vigilance. Future lab-based and observational research is needed to explore this tentative explanation.

Second, the lack of behavioral findings raises the question of how online vigilance manifests behaviorally if not via the theorized means. Initially, it may be that the smartphone features that we developed are not good representations of online vigilant behavior, and that we ought to have operationalized this construct differently. An exploration of alternative operationalizations, however, did not immediately lead to markedly different results. Additionally, perhaps smartphone behavior alone is not indicative enough, and we should also glean other behavioral data, such as people’s computer use. After all, (work) e-mails may be something adults are highly online vigilant over, but these e-mails might be dealt with on the computer more so than the smartphone. Also, many popular communication channels are simultaneously available on smartphones and other devices, and there may be a ‘handover’ from one device to the other depending on the context one is in. Hence, to develop valid behavioral features of online vigilance, we might have to look at all digital media behaviors.

Finally, one crucial conclusion that we might draw from our collective evidence is that subjective experience trumps behavior in explaining the downstream implications of digital media use. This perspective aligns with recent findings by media scholars [such as Ernala et al. (2022); Lee et al. (2021); Lee & Hancock, 2023; M. M. P. Vanden Abeele (2021)]. Their work underlines the significance of subjective beliefs, experiences and mindsets in relation to digital connectivity in explaining the impact of media use on different psychological outcomes. Online vigilance might, similarly, represent such a mindset or media experience which is better captured subjectively, rather than through the use of different behavioral features.

This study is not without limitations, of which the first is undoubtedly the use of a self-selected sample. By cooperating with a local newspaper, advertisements for the study were solely published on their website and in print, to an audience that is, on average, older and more highly educated. Individuals also participated voluntarily, which likely indicates their interest in the topic. This approach has likely narrowed down our potential research population and may have potentially led to certain biases. Alternatively, intensive longitudinal studies on online vigilance in adult (rather than in student) populations are generally less prevalent, and the fact that participants in our study were not financially incentivized for their participation has likely improved data quality overall.

We wish to remark here that extending this line of research to a clinical population, in particular, could provide insights into potential implications of online vigilance on mental health and wellbeing in vulnerable groups. Such research could also consider the role of accumulated fatigue in the relationship between online vigilance and well-being, allowing a more nuanced understanding of its long-term impacts, tipping points, and potential resilience factors.

A second major limitation of this study is that participation in this study required smartphone ownership and use, for the completion of daily ESM questionnaires. This has two major ramifications. First, during the recruitment process, the research team received a number of requests for participation from adults who did not own or use a smartphone. Because managing the data collection was already extremely challenging, we did not act upon these requests, but it is important to remark that this subpopulation who deliberately chooses not to use a (smart) mobile phone may provide crucial insight into digital well-being phenomena. Second, in the end-of-study questionnaire, a number of individuals pointed towards the irony of using a smartphone for a research project on the smartphone and indicated that their 'missing values' were often due to deliberate decisions to go offline. Some participants also noted being vigilant for the study notifications, causing an obvious confound. Here we lay bare an intrinsic methodological problem of ESM research for the study of digital well-being and digital disconnection, an observation also made by Klingelhoefer et al. (2023), namely that you intervene on the very reality that you aim to capture, thus refuting the claim of ecological validity of ESM research.

Third, in our study we observed a high correlation between availability pressure and online vigilance, suggesting a notable conceptual overlap that warrants further consideration. This overlap may imply that in certain contexts, the perceived pressure to be available online and the resultant vigilance in monitoring digital communications are so closely intertwined that they might influence each other more profoundly than initially anticipated. This finding raises important questions about the independence of these constructs and the potential for a shared underlying factor. It also suggests that the measures used may capture elements of both constructs, thus potentially affecting the clarity of the distinctions drawn between them. Future research should explore this further, potentially by exploring through lab experimentation if a meaningful distinction can be drawn.

Finally, while our study explored lagged relationships to address causality concerns, reverse causality as well as the presence of unmeasured third variables influencing the observed outcomes need further exploration. It is plausible, for instance, that participants were more vigilant to their online world during periods of elevated stress, and that this stress contributed to increased mental fatigue.

Notwithstanding these limitations, our study also has several strengths: Our findings draw from a very large sample (1,315 participants, utilizing around 42,000 datapoints to test contemporaneous and lagged models that control for mental fatigue's autoregressive effect), that included a large number of workers, with ages spanning from 18 to 82 years old, and showed effects that were directionally consistent across models. Thus, we have some confidence that the effects in our study are unlikely to reflect type 1 errors, and that the identified effects can tell us something about what excessive online vigilance may lead to in populations particularly prone to experiencing burnout, namely, those who are actively employed. Moreover, a positive link between online vigilance and mental fatigue was established both after controlling for likely confounds (e.g., autoregression effects) and after temporally separating online vigilance from mental fatigue, thus offering some suggestion of potential causality. Our use of experience sampling also reduced concerns about recall bias, a critical issue in some previous research that examined the link between online vigilance and outcomes conceptually related to mental fatigue (Freytag et al., 2021; Reinecke et al., 2018), whereby participants had to recall their experiences from many more hours prior. Lastly, our work adds to extant scholarship work on online vigilance and adjacent outcomes such as stress (Freytag et al., 2021; Gilbert et al., 2023; Reinecke et al., 2018) or mind-wandering and mindfulness (Johannes et al., 2018). As such, it further showcases the importance of the subjective experience of being 'always on' impacting well-being. To summarize, then, putting aside the necessary caution one must have for a first test of effects, the strength of the present findings edges forward the idea that increased levels of online vigilance are likely to elevate mental fatigue, and further down the line, if unchecked, the broader symptoms of burnout.

Always On, Always Rushed for Time?

Exploring Momentary Associations Between Passively Sensed Smartphone Use, Feeling Rushed, and Perceived Task Juggling[10]

[10]Published as: Van Gaeveren, K., Murphy, S., De Segovia Vicente, D., & Vanden Abeele, M. (2025). Always on, always rushed for time? Exploring momentary associations between passively sensed smartphone use, feeling rushed, and perceived task juggling. *Mobile Media & Communication*. https://doi.org/10.1177/20501579251377010

Abstract

This mixed-methods study combines experience sampling and smartphone log data to explore momentary associations between passively sensed smartphone use, feeling rushed, and perceived task juggling in daily life. Using data from 774 adults, we analyze how four use features (frequency, duration, fragmentation and notifications received) of four mobile app categories (email, social media, chat, and work communication) affect how rushed people feel, both directly and indirectly by increasing their perceived juggling load. At the between-person level, persons who use work communication app more frequently and longer, and who receive more chat notifications, also felt more rushed on average. Findings also revealed within-person associations in the theorized direction (e.g., increased email frequency predicting increased task load and feelings of being rushed) for nearly all features of three out of four examined app categories, with social media use features as a noticeable exception. Perceived task juggling mediated these associations, suggesting that the examined smartphone use features contribute to feeling rushed by increasing the (real or perceived) load of tasks that people juggle within and across role domains. Some patterns differed based on age, gender, parenthood and segmentation preference. Taken together, these findings support the theoretical link between technology use and experiences of time scarcity.

Keywords: Mobile communication, feeling rushed, time pressure, acceleration theory, juggling load, log data, experience sampling, passive sensing, social media, email, messaging, chat, work communication

Introduction

The advent of mobile communication has deeply impacted everyday life, promoting a culture of 24/7 connectivity in which people are 'always on' (Castells, 2007; M. M. P. Vanden Abeele, 2021; Vorderer et al., 2017). This always-on lifestyle is linked to perceptions of time. In line with Rosa's (2017) acceleration theory, for instance, which posits that technological advancements increase the perceived pace of life, mobile connectivity is theorized to play a role in why people generally feel rushed, or as Wajcman (2008) puts it, as if they are living life 'in the fast lane'.

It seems to matter *what* smartphones are used for, however: Ross et al. (2023), for instance, found social app use was associated with *de*celerated time perception, whereas messaging app use was associated with *ac*celerated time passage among more intense users of such apps. An explanation for this differential pattern of results can be found in Bittman et al.'s (2009) study on the link between mobile communication and experience of feeling rushed. The findings from this study suggest that, especially in workplace contexts, mobile communication in the form of calling and texting has the capacity to 'compress time' by increasing task density and evoking more frequent task switching. As Ross et al. (2023) note, social media use, on the other hand, might rather signal 'downtime', i.e. moments in which time is less compressed than usual.

While theoretical work on mobile communication and time perception is extensive, empirical studies using behavioral data remain scarce. Hence, a first aim of this study is to examine momentary associations between passively sensed mobile media use and feeling rushed for time in a large sample of adults (N = 774). We hereby explore the role of four different app categories, namely email (e.g. Outlook), chat (e.g., WhatsApp), social media (e.g., Instagram) and work communication (e.g., Teams, Slack), and examine theoretically relevant behavioral features of their use that go beyond frequency and duration, also exploring their *fragmentation* and the number of *notifications* received for them.

Second, the mechanisms linking mobile communication to feeling rushed remain understudied. As mentioned above, the work of Bittman et al. (2009) and others (Kalman et al., 2021; Lanctot & Duxbury, 2021; Pritchard & Symon, 2023; Thulin & Vilhelmson, 2022) suggests that mobile communication can lead to an *intensification* of work, with multiple tasks being squeezed and multitasked. Following K. J. Williams et al. (1991), we define this experience as 'task juggling'. Psychological research suggests task juggling elevates cognitive load, fragments attention and heightens subjective feelings of urgency, leading to a more fragmented perception of time (Baethge & and Rigotti, 2013; Mark et al., 2008; Sussman & Sekuler, 2022). Moreover, because of its boundary blurring potential, mobile communication may not only foster the juggling of tasks within a social role, but also across roles, for instance when individuals juggle both contractual work and unpaid household tasks in the same timespan (Lanctot & Duxbury, 2021; Nagy et al., 2021; Wajcman, 2015). A second aim of this study is to examine if, at the momentary level, the experience of intra- and interrole juggling explains the association between mobile communication and feeling rushed.

Finally, scholars such as Wajcman (2008) indicate factors such as gender and parenthood likely moderate the aforementioned associations. Moreover, mobile media (e.g., Ross et al., 2023; Vaid et al., 2024) and media effects scholars (e.g., Valkenburg et al., 2024) increasingly pay attention to effect heterogeneity. A third aim is therefore to explore this heterogeneity, and how it relates to sociodemographic factors such as gender, age, parenthood and employment status.

To address these aims we conducted a mixed-methods study involving 774 adult participants, from whom we collected both experience sampling (ESM) data and passively sensed smartphone use data. Before providing further details on the methodology, however, we first explain our theoretical rationale.

Theoretical Framework

Acceleration and Mobile Communication

Mobile communication is regarded as a contributing force to people's harried experience of time (Bittman et al., 2009; Ling, 2017). At first glance, this may be counterintuitive. After all, mobile connectivity offers its users more flexibility to organize their work and personal lives, which could supposedly help them to free up time. But research shows mobile connectivity has rather led to increased productivity demands (Prommer, 2019; Santarius & Bergener, 2020). Indeed, as a result of mobile communication even the interstices of life, i.e. the moments in-between the "spheres of home, work and friends" (Görland, 2019, p. 323) are now filled up with social activities and work (Santarius & Bergener, 2020).

In view of the above, it might not surprise that studies link mobile connectivity to perceptions of an overall increase in the pace of life. Thulin & Vilhelmson (2022), for example, characterize mobile technologies as the "pacesetters of everyday rhythms of work" (p. 262) as they find that for many, these devices, and in particular mobile email applications, are used to synchronize tasks and updates with colleagues, leading to expectations of higher availability and an increase in overall work pace. Similarly, Bittman et al. (2009) found positive associations between mobile phone use, work intensification, and time pressure, albeit only for men. From a theoretical point of view, these observations are illustrative for a society that is accelerating (Rosa, 2017): the rapid technological change in mobile communication has contributed to social change, i.e. in the form of the emergence of a culture of 24/7 connectivity, in which individuals increasingly perceive time as fast-paced, and experience a constant feeling of being rushed.

Work by Mullan & Wajcman (2019) nuances the relationship between mobile use and subjective time pressure, however: They found that, while the total time spent performing work on mobile phones had no unique effect on time pressure, short "micro-episodes" of quick checking or monitoring of incoming communications already influence how rushed people feel. This highlights the importance of considering the nature of mobile interruptions (Wajcman, 2008).

The Nature of Mobile Communication: Frequency, Duration, Fragmentation and Notifications of Different App Categories

To understand the *nature* of mobile communication, we should not treat 'smartphone screen time' as a monolithic concept, but rather consider its theoretically relevant facets (Kaye et al., 2020; Meier & Reinecke, 2021). To situate these theoretically relevant facets, we can draw from Meier and Reinecke's (2021) taxonomy of computer-mediated communication (CMC), which distinguishes different analytical levels of mobile communication.

One such facet is what *category (or type) of mobile application* the phone is used for. These categories often share distinct features (Meier & Reinecke, 2021). Applying this framework, our study specifically investigates four theoretically relevant mobile application categories (email, work communication, messaging and social media) and explores four of their usage features (frequency, duration, notifications and fragmentation). This choice is based on prior research indicating meaningful variation based on both app categories and usage characteristics (e.g. notifications or fragmented use patterns) related to stress and distraction (Aalbers et al., 2023; Siebers et al., 2024).

Work-related apps (e.g., Teams, Slack) might increase feelings of being rushed more than mobile messaging or social media apps, as this category can be linked to work intensification (Lanctot & Duxbury, 2021; Mullan & Wajcman, 2019; Thulin et al., 2019; Thulin & Vilhelmson, 2022; Wajcman, 2015). With respect to usage characteristics, there are also time-relevant features of mobile communication that go beyond screentime and duration: First, mobile communication typically comes with *notification systems*. Prior research already shows that receiving numerous notifications can be overwhelming and stressful (Lanctot & Duxbury, 2021; Ytre-Arne et al., 2020). We build further on this observation, hypothesizing that receiving more notifications — especially from mobile email or work communication — can create situations of high urgency, which may make people feel more rushed.

Second, mobile communication typically comes with *fragmentation*, which refers to frequent, short bursts of engagement with the device dispersed across time (Siebers et al., 2024). These may play a role in experiences of harriedness (Wajcman, 2008). In the current study, we include a fragmentation feature based on the work by Siebers et al. (2024), who have shown that the fragmented nature of interacting with a mobile device is associated with distraction.

Based on the above-mentioned evidence of mobile technology use as a source of harriedness, our first research aim is to explore how different facets of mobile media use associate with feeling rushed, where we hypothesize that:

> **H1:** At the within-person level, individuals will feel more rushed when they have used mobile communication apps (i.e., mobile email, mobile work communication, mobile messaging and mobile social media use) more frequently (H1a), for a longer duration of time (H1b), received a larger number of notifications for these app activities (H1c) and have engaged in these activities in a more fragmented nature (H1d).
>
> **H2:** We expect the above pattern of associations to be stronger for mobile email and mobile work communication app use than for mobile messaging and mobile social media use.

Task Juggling as a Mediation Process

While mobile devices may *directly* impact how rushed we feel, we also need to consider the processes of *how* these effects come to be. One such process is likely the experience of "juggling" multiple tasks, within and across role domains. Mobile devices allow for anytime, anyplace connectivity, and hence for communication about new tasks and responsibilities, both within and outside the current social role one is performing (Duxbury et al., 2014; M. M. P. Vanden Abeele et al., 2018). This leads to mobile connectivity having a complex and even paradoxical relationship with autonomy (M. M. P. Vanden Abeele, 2021): While we gain autonomy by switching more flexibly between tasks pertaining to our various social roles, the inescapable flipside of this flexibility is that we are also constantly exposed to the potential of incoming demands of these roles, that therefore can more easily blur and come into conflict, leading to (role) pressure and stress.

Prior research already shows that being interrupted by, and needing to manage, overlapping roles and tasks can lead to negative well-being outcomes, such as exhaustion (Derks et al., 2021) and stress (Cornwell, 2013). We argue here that another outcome may be feeling more rushed. Indeed, an accelerated experience of time may stem from the sentiment that we are constantly 'juggling' or 'multi-tasking' (potentially conflicting) tasks, both within and across different role domains (Wajcman, 2008), as this may lead to our work in these domains being intensified (Bittman et al., 2009). A recent diary study (Lu, 2024), for instance, has shown that an increase in work-time fragmentation, characterized by more disruptions in the workday schedule, impacts the amount of time pressure experienced. This is presumably because the necessity of frequent role switching to meet various role demands contributes to this pressure (Cornwell, 2013): These different roles vie for our time and attention, often leading to role conflicts and role blurring which can make us feel harried or rushed for time (Lu, 2024). Given the ubiquity of mobile devices and the communication they facilitate, multiple disruptions throughout the day seem inevitable, potentially triggering a higher task load within and across our work and private roles.

Psychological literature on time perception provides additional insights into how task juggling, whether induced by mobile communication, might intensify feelings of being rushed. Specifically, some of the mechanisms through which juggling might mediate this relationship include task density (the number of tasks packed into a limited timeframe), frequent task-switching (shifting attention between distinct activities), interruptions needing reprioritization of ongoing tasks, and the general disruption of task flow (Baethge & and Rigotti, 2013; Steyvers et al., 2019; Wearden, 2016; Wittmann, 2009). Each of these mechanisms impose a cognitive load that can fragment attention and can contribute to subjective experiences of increased urgency or time pressure. In addition, states of high cognitive load are likely to lead to increased arousal, and the interplay between these states can further intensify perceptions of accelerated time (Wittmann, 2009). Given that mobile communication frequently involves these types of attentional switching and disruptions to task prioritization and workflows, it likely increases experiences of feeling rushed.

Therefore, as second research aim, we test if juggling load mediates the effect of mobile media use on feeling rushed:

> **H3:** At the within-person level, perceived juggling load mediates the association between mobile communication and feeling rushed: At moments when individuals engage more in mobile communication (as expressed in frequency, duration, number of notifications, and fragmentation), they report a greater juggling load, and this load in turn explains momentary feelings of being rushed.

Person- and Situation-Specificity

Finally, we may expect mobile communication to impact experiences of *feeling rushed* and *juggling load* differently for different individuals, depending on both the context and personal characteristics of these individuals. Vostal (2021) recently highlighted the importance of considering the dynamic and person-specific experiences of acceleration. They argue for a view where individuals have “temporal agency” and thus the capacity to develop strategies to navigate acceleration experiences, which can differ from moment to moment. Individuals may integrate their mobile devices differently as part of these strategies. Yet, as Vostal (2021) notes, not every person may have equal access to every strategy, and not every strategy is likely to be equally effective for every person.

Structural factors such as preferences for role integration and segmentation, employment status, household responsibilities and gender can shape how individuals experience mobile communication’s impact (Grotto & Mills, 2023; Piszczek, 2017; Thulin et al., 2019; Wajcman, 2015). Grotto & Mills (2023), for instance, showed that illegitimate interruptions had a larger effect on well-being for men compared to women, possibly due to our social expectations on gender roles. Work by Lu (2024) has also shown that the fragmentation of work time can have a different impact on the amount of time pressure experienced depending on gender and household responsibilities. They found that women without children showed a stronger relationship between work time fragmentation and time pressure compared to mothers, and, surprisingly, the opposite relation for fathers. Finally, Duxbury et al. (2017) found that women's feelings of being time-pressed and overwhelmed by role demands were more strongly influenced by their home experiences compared to men.

These findings showcase the importance of taking dynamic and person-specific effects into account — a claim which has recently been voiced by many mobile media scholars (e.g., Vaid et al., 2024; Valkenburg et al., 2024; M. M. P. Vanden Abeele, 2021). As such, the third aim of this study is to investigate the dynamic and person-specific nature of the relationship between mobile media use and time experiences. To that end, we explore:

> **RQ1:** To what extent are person-specific associations observable between mobile communication, juggling load and feeling rushed?
>
> **RQ2:** What differences can be observed in these associations based on gender, age, parenthood and segmentation preferences?

Methods

Sample

This study draws from a large-scale citizen science project (approved by the IRB of Ghent University, code: 2022/11). A total of 1,315 participants provided consent prior to completing two weeks of mobile experience sampling, with data collection starting in November 2022. From 916 Android[11] users we collected over 7 million unique smartphone trace events.

After data cleaning, the final dataset included 774 participants with sufficient ESM and smartphone data. The average age within this sample was 38.32 years (*SD* = 11.17), with participants identifying predominantly as female, actively working and having higher education qualifications.

Procedure

Participants installed apps for questionnaires (m-Path) and smartphone monitoring (MobileDNAPlus).

Over two weeks, participants received six daily questionnaires (7:30 a.m. – 10:45 p.m.) assessing feelings of being rushed and juggling load, among other topics outside this study's scope. Our OSF page (see below) gives an overview of all items included.

Furthermore, during this two-week period, our passive monitoring tool captured participant's app usage (timestamped applications), sessions (timestamped screen on/off events) and notifications (timestamped notifications received and interacted with). These trace data were then aggregated into different features (see Behavioral measures below), matching the ESM time windows.

Measures

4.4.c.a Momentary self-report measures

Feeling rushed was measured by asking participants: "Since getting up this morning/Since the last beep… I felt rushed for time.", with the following options: (1) Not at all, (2) Very rarely, (3)

[11] iPhone trace data could not be collected due to technical limitations of the operating system.

Rarely, (4) Sometimes, (5) Often, (6) Very often (7) All the time. This measure, adapted from Robinson (2013), incorporated a bipolar response scale.

The **Juggling load** measure, adapted from K. J. Williams et al. (1991), asked participants whether they switched tasks ('no', 'between non-work tasks', 'between work tasks', 'between both work and non-work tasks'). We recoded responses ordinally: (1) no juggling, (2) within-domain juggling (work or private), and (3) cross-domain juggling. This decision was made given we theorized that while some task juggling was likely to manifest as an increased task load and increased feelings of being rushed relative to no task juggling, it seemed plausible that juggling tasks across domains would constitute *additional* load. For instance, completing tasks related to one's personal life while at work in the eyes of most employers is not a justifiable use of one's work time, thus receiving communication to complete personal life tasks may take the form of adding the task to the many work tasks one already must engage with. In contrast, receiving a work task via email when at work may allow one to justify reprioritizing one's work load by, for instance, delaying some tasks until the following day. We also consider it plausible that across-domain juggling will generally be more disruptive (e.g., to one's workflow) than within-domain juggling, further justifying our operationalization.

4.4.c.b Behavioral measures

Based on the smartphone trace data, we created the following measures, aggregated per application category and per self-report time window: **Duration** sums the total screentime for each of the following four app categories: chat (e.g. WhatsApp), social (e.g. Facebook), email (e.g. Outlook) and work communication (e.g. Microsoft Teams). **Frequency** counts the total amount of app launches for each of these same four categories. **Fragmentation** looks at the dispersion of these two dimensions across time: A low value signifies that less time was spent on this app category and this across fewer sessions, while a high value indicates more time spent, and this in frequent and short bursts (see Siebers et al., 2024 for further information). Finally, **notifications** sums the number of incoming notifications for the four categories of app activities. Operationally, apps were grouped into categories based on their default categorization from the Google Play Store, combined with manual validation. Our OSF page includes the full list of apps and their categorization.

4.4.c.c Socio-demographic, trait and control variables

We included the following sociodemographic and trait self-report measures: **Gender**, which was recoded as (0) Male, (1) Female. **Age**, participant age in years. **Parenthood status**, recoded as (0) no children, (1) one child or more. **Segmentation preference**, a Likert-like scale stating "I prefer to keep my work life and private life separate" with (1) completely disagree to (5) completely agree. The variable was then dichotomized, with scores below four coded as (0), while scores of four and five coded as (1). In addition, we included two control variables: **Work activity** (binary indicator: 1 = working, 0 = not working) and **Time of day**, based on the ESM answer timestamp, dummy-coded into three variables, morning (6h – 12h), afternoon (12h – 18h) and evening (18h – 24h), with afternoon taken as the reference category.

Data Analysis

Data preparation and descriptive statistics were conducted in Python using the Pandas package, with the repeated measures correlations obtained through the pingouin package. Level-1 Mediation analyses were conducted in R using the Lavaan package, while mixed-effects and person-specific models were conducted using the lme4 package. Linear mixed-effects models examined associations between rushed and smartphone use. For mediation and person-specific analyses, all variables were measured at the same time-point, and models included the binary *work activity* and *time of day* control variables. Cross-level interaction models did not include these binary controls. All Lavaan models showed adequate fit (*CFI*: above .90; *RMSEA*: lower than .08; *SRMR*: lower than .05) according to current standards (Marsh et al., 2004). Lastly, we employed two-tailed testing and considered effects significant if p-values were .05 or less.

Transparency and Openness

All code, processed data and materials for reproducibility can be found on OSF (https://osf.io/ztxmh).

Results

Descriptives

Table 4.1 presents within-person descriptives, showing that participants rarely felt rushed, but indicated feeling somewhat (4) to very (7) rushed and to be juggling in 39% of the questionnaires. Juggling load was significantly correlated with feeling rushed at both the within- and between-person levels. Working was also significantly correlated with juggling load and feeling rushed at both levels. At the between-person level, feeling rushed correlated with gender (higher among women: r = .12), age (r = -.15), and segmentation preference (r = .09) while juggling load correlated with parenthood (r = .27) and segmentation preference (r = .14).

Table 4.2 reports the *daily* averages of mobile communication measures, for participants with more than seven days of log data, but note that for the further analyses the full sample was considered. Participants' average daily smartphone use was 2hr and 25min with considerable variation across application categories.

Table 4.1: Descriptives for each within-person level measurement (Level 1: per ESM time window).

	Mean	SD	Min	Max
rushed	2.91	1.67	1.0	7.0
juggling load	0.49	0.67	0.0	2.0
chat duration	2.73	8.46	0.0	267.67
chat frequency	5.74	11.24	0.0	273.0
chat notifications	3.64	7.77	0.0	260.0
chat fragmentation	0.25	1.04	0.0	42.84
social duration	2.96	8.36	0.0	201.1
social frequency	2.93	8.31	0.0	279.0
social notifications	0.61	2.43	0.0	131.0
social fragmentation	0.27	1.23	0.0	75.54
email duration	0.80	2.75	0.0	98.25
email frequency	2.55	6.06	0.0	155.0
email notifications	1.05	3.80	0.0	207.0
email fragmentation	0.07	0.35	0.0	27.51
w. comm.[a] duration	0.05	0.66	0.0	41.14
w. comm. frequency	0.10	0.94	0.0	36.0
w. comm. notifications	0.12	1.13	0.0	72.0
w. comm. fragmentation	0.00	0.07	0.0	6.27

N = 42,858. [a] w. comm. = work communication. All duration variables are expressed in minutes *per questionnaire time window. The average duration of a time window was 3h 17m 43s. Descriptives we calculated across all data points, without first aggregating per participant.*

Table 4.2: Daily smartphone descriptives for participants with > 6 days of log data.

	N	Mean	SD	Min	Max
total duration	568	145.50	66.79	22.01	433.90
total frequency	568	90.98	45.54	10.22	284.82
total notifications	612	26.81	22.26	1.50	225.17
chat duration	567	18.10	17.58	0.30	131.02
chat frequency	567	18.98	13.15	1.60	108.91
chat notifications	606	16.37	13.65	1.00	109.62
social duration	488	23.47	23.10	0.00	193.99
social frequency	488	12.18	12.39	1.00	95.45
social notifications	383	5.57	9.50	1.00	109.08
email duration	548	5.64	6.33	0.03	61.02
email frequency	548	9.79	8.75	1.00	58.55
email notifications	402	8.52	10.04	1.00	86.25
w. comm. duration	162	2.26	2.39	0.00	14.28
w. comm. frequency	162	2.90	2.46	1.00	23.67
w. comm. notifications	135	4.97	5.16	1.00	34.71

Daily smartphone descriptives for participants with > 6 days of log data.

Calculated as the average amount per participant, per day. All duration variables are expressed in minutes *per day.*

Smartphone Use and Feeling Rushed

Table 4.3 shows the within- and between-person correlations. We found partial support: Social media use was not significantly correlated with feeling rushed, but for the other three app categories all features (frequency, duration, fragmentation and notifications) were positive predictors.

Overall, correlations were small (r = .03 to .07). Notifications from email (r = .07) and work apps (r = .06) were most strongly related to feeling rushed, while fragmented use had weaker associations.

At the between-person level, several of the usage indicators related to work communication, email, social media and chat apps showed significant positive correlations with feeling rushed, again offering partial support.

Our second hypothesis expected the associations of email and work communication apps to be stronger than those for social and chat apps. Figure 4.1 shows the coefficient estimates and confidence intervals (CI) of sixteen (four features across four categories) mixed-effects models, and provides partial support: Examining each category's CI, we saw that *social media* features showed no overlap with *email* or *work communication* features, however, *chat* features did.

Table 4.3: Within- (lower diagonal) and between-person (upper diagonal) correlations.

	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20
1. rushed		.50***	.27***	.12***	.07	.10**	.15***	.09*	.02	.08*	.07*	.03	.05	.07*	.07*	.05	.11**	.11**	.09**	.10**
2. juggling load	.38***		.38***	.05	.02	.09*	.08*	.08*	.01	.04	−.02	.07*	.12**	.14***	.03	.16***	.08*	.12**	.12**	.07*
3. work [a]	.32***	.45***		−.14***	−.02	.06	.10**	.05	.01	.04	.03	.06	.02	.04	.05	.07*	.14***	.15***	.16***	.11**
4. gender [b]	−.00	.00	.02**		.14***	.11**	.04	.09*	.15***	.11**	.03	.08*	−.01	.00	−.03	−.03	−.05	−.03	−.06	−.03
5. chat duration	.03***	.05***	.04***	−.01*		.52***	.53***	.82***	.31***	.26***	.16***	.28***	.15***	.14***	.10**	.19***	−.00	−.02	−.04	.01
6. chat frequency	.03***	.06***	.05***	−.01	.49***		.54***	.66***	.37***	.55***	.25***	.49***	.19***	.35***	.07*	.31***	.01	.07	.02	.03
7. chat notifications	.06***	.09***	.09***	.00	.33***	.47***		.60***	.33***	.41***	.33***	.37***	.03	.07*	.23***	.12***	−.02	−.01	−.02	−.01
8. chat fragm.	.04***	.05***	.05***	−.01	.74***	.58***	.38***		.35***	.45***	.25***	.54***	.19***	.23***	.09*	.40***	−.00	−.01	−.03	.02
9. social duration	−.01	−.00	−.03***	−.01	.15***	.27***	.16***	.19***		.67***	.39***	.83***	.10**	.18***	.12**	.18***	.01	.04	.01	.02
10. social freq.	−.01	.01**	−.00	−.00	.19***	.42***	.24***	.27***	.63***		.53***	.72***	.05	.23***	.05	.18***	−.00	.05	.01	.01
11. social notif.	.01	.03***	.03***	.00	.11***	.19***	.28***	.14***	.22***	.37***		.42***	−.02	.01	.13***	.05	−.01	.00	.02	−.00
12. social fragm.	.01	.03***	.01**	−.00	.17***	.33***	.20***	.36***	.71***	.56***	.23***		.13***	.24***	.08*	.36***	−.01	.02	.00	.02
13. email duration	.04***	.05***	.05***	.00	.13***	.22***	.11***	.15***	.12***	.13***	.05***	.11***		.71***	.04	.84***	.11**	.12***	.09*	.12***
14. email frequency	.05***	.06***	.07***	−.00	.20***	.40***	.19***	.24***	.23***	.28***	.09***	.24***	.64***		.04	.71***	.13***	.18***	.13***	.14***
15. email notif.	.07***	.09***	.13***	.01	.11***	.15***	.24***	.15***	.08***	.09***	.14***	.10***	.14***	.18***		.06	.07	.09*	.12***	.09*
16. email fragm.	.04***	.05***	.06***	−.00	.14***	.28***	.17***	.27***	.12***	.15***	.06***	.21***	.75***	.60***	.19***		.11**	.12**	.10**	.15***
17. w. comm. dur.	.03***	.04***	.05***	.00	.02***	.04***	.03***	.02***	.02***	.02***	.01*	.02***	.06***	.08***	.07***	.08***		.87***	.65***	.93***
18. w. comm. freq.	.04***	.05***	.08***	.01	.03***	.07***	.04***	.03***	.05***	.06***	.03***	.05***	.09***	.14***	.10***	.10***	.58***		.67***	.84***
19. w. comm. notif.	.06***	.07***	.11***	.00	.01**	.03***	.05***	.01*	.02**	.01**	.03***	.02***	.03***	.04***	.18***	.04***	.19***	.34***		.60***
20. w. comm. frag.	.03***	.04***	.05***	.00	.03***	.05***	.04***	.04***	.03***	.03***	.02***	.06***	.10***	.10***	.13***	.16***	.81***	.58***	.21***	

[a] work: 0 = no work activity reported, 1 = work activity reported.

[b] gender: 0 = male, 1 = female. * $p < .05$ ** $p < .01$ *** $p < .001$

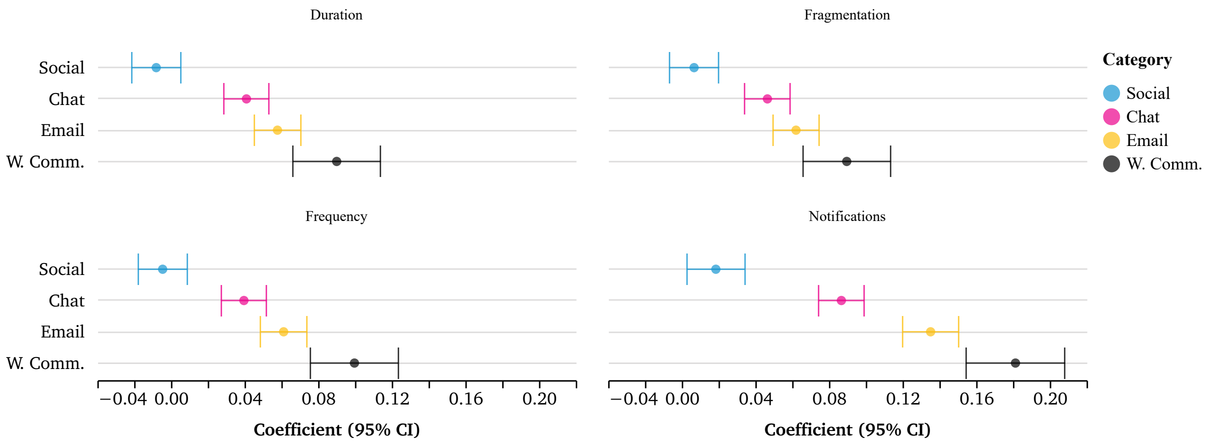


Figure 4.1: Coefficients and Confidence Intervals per Feature and Category on Feeling Rushed

Note: *This figure shows the effects of the standardized coefficients on the original (non-standardized) rushed variable.* W. Comm. *is short for the "work communication" category.*

Juggling Load Mediation

Thirdly, we hypothesized that the link between mobile media use and feeling rushed would be partly explained by task juggling, such that mobile communication promotes a greater juggling load, which in turn leads to feeling more rushed. The within-person correlations table (see Table 4.3) already indicates that, overall, associations between mobile communication and juggling load appeared to be slightly stronger than those between mobile communication and feeling rushed. To formally test this hypothesis, however, we performed a series of mediation analyses, with each model including a single behavioral representation of mobile media use. Hence, in total, we ran sixteen models (four features across four categories).

In addition, we controlled for momentary work engagement due to its link with time pressure (Mullan & Wajcman, 2019; Pritchard & Symon, 2023), and for the time of day, as this can influence both subjective time experiences and mobile communication usage patterns (e.g., Ross et al., 2023).

Figure 4.2 shows the diagram for each of these models, grouped per category. Table 8.1 in the appendix contains the full output. We report findings per mobile communication activity.

For *mobile messaging or chatting*, we see that all features showed a small effect on feeling rushed through role juggling. Chat interactions indirectly increased feelings of being rushed through juggling load, while direct effects were suppressed when controlling for work and time of day.

For *social media* use, the pattern was different: When controlling for juggling load, interacting with social media failed to show any direct or total effects, while there were only small, but significant and positive indirect effects for social media frequency, duration and fragmentation on juggling load.

All *email* features directly and indirectly increased feelings of being rushed. When investigating the path through juggling load, all features showed a significant *total* effect, with the amount of email notification received being the strongest predictor (β = .035).

Finally, the notifications received from *work communication* apps showed a direct effect on feeling rushed, while, once more, when looking at the indirect effects through juggling load, we saw all four features showing a significant effect. Similar to the email category, the models testing duration and frequency of work communication use showed smaller total effect sizes (β = .009, .008) compared to the model with notifications received (β = .017).

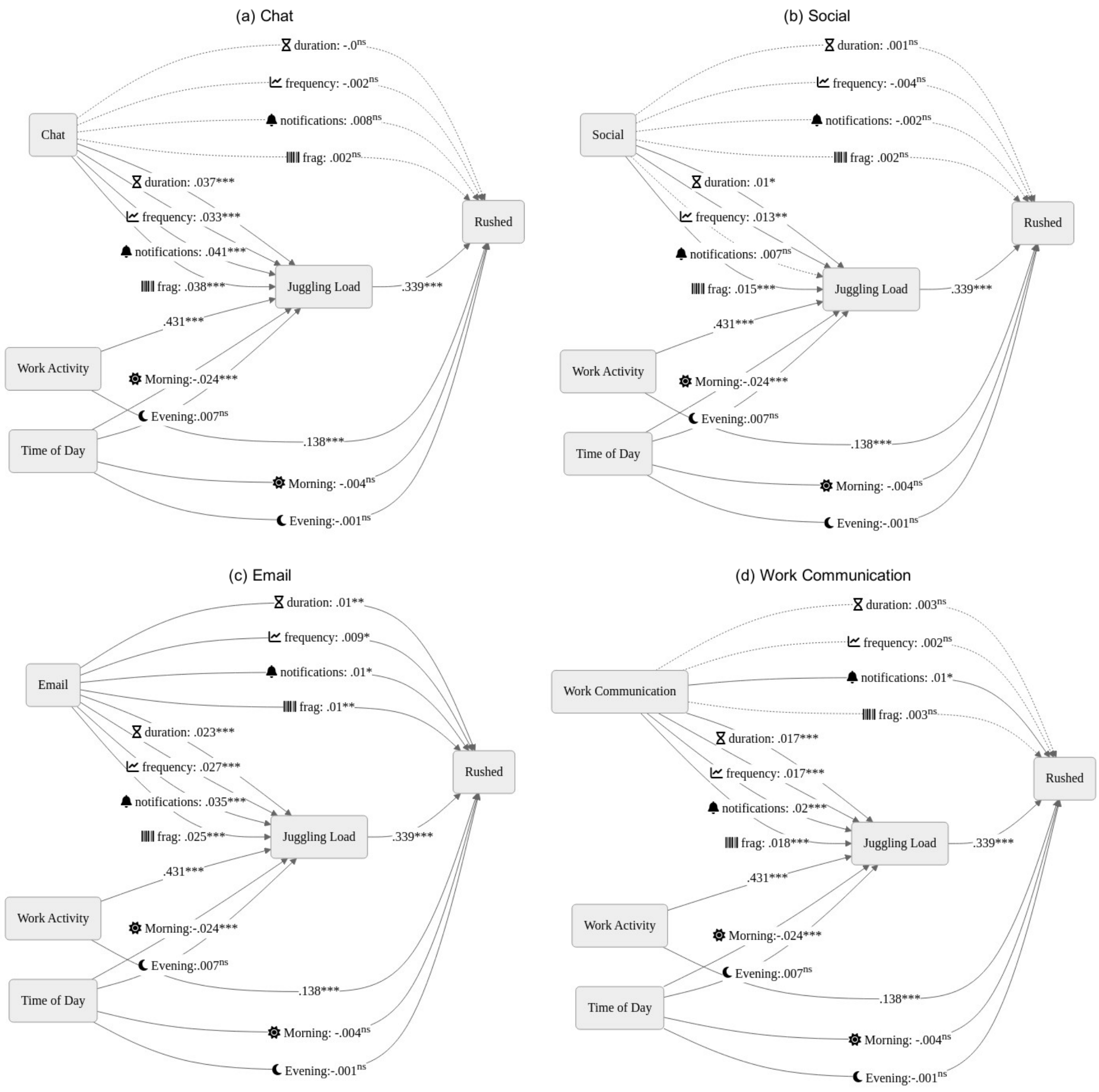


Figure 4.2: Mediation Models

Person-specific Effects

As a third and final aim, we explored the person-specificity of the associations between mobile communication and feeling rushed and juggling load. Table 8.2 and Table 8.3 in the appendix show the distribution of person-specific effects on *feeling rushed*, and *juggling load*, respectively, again controlling for having worked and time of day. Most participants showed negligible associations (coefficients < .05), across all our features, with many effects being non-significant.

For feeling rushed, chat and email *notifications received* showed both positive and negative subgroups, with heterogeneity intervals showing associations nearly three to six times as intense compared to their average associations.

Juggling load, in contrast, showed less variability for chat notifications received, but a similar pattern for *email notifications*. Here, the heterogeneity interval indicated that some participants showed negative associations nearly three times as intense, and some positive and five times as intense than the average.

While it is possible that participants experienced certain media effects very differently, these results should be interpreted with caution. Many of these effects were likely influenced substantially by the lower amount of datapoints available per participant. They may very likely reflect random variation of a true null effect, rather than opposite effects canceling each other out (Johannes, 2020). Nonetheless, cross-level interactions revealed that sociodemographic factors explained some of this heterogeneity (see Table 8.4). Older participants had weaker associations between chat notifications and feeling rushed or juggling load. In addition, the association between juggling load and both chat and email notifications received was stronger for participants preferring segmentation. Finally, participants with children tended to show stronger associations between email notifications received and feeling rushed.

Discussion

The first two aims of this study were to examine (1) whether mobile communication is associated with how rushed people feel, and to explore (2) if task juggling serves as a potential mediating process explaining such an association. The third aim was to explore whether there is heterogeneity in these associations, and if so, if that heterogeneity can be explained by person- and context-specific factors. To address these aims, we drew from a dataset of 42,861 data points combining passively sensed smartphone data with multiple daily self-reports and survey data gathered from 774 adult participants.

With respect to the first aim, this study found that various features of the four examined mobile communication app categories were positively correlated with feeling rushed. In other words, fluctuations in chat, email, work communication app use were positively related to feeling rushed. This idea is also supported by our between-person level analyses, which indicate that individuals who in general interacted more with work communication apps and received more notifications by chat apps reported higher overall levels of feeling rushed. The findings also provide support for our second aim, as perceived juggling load mediated the relationship between several features of mobile communication and feeling rushed.

Overall, these findings align both with acceleration theory (Rosa, 2017) and psychological theories of time perception (Baethge & and Rigotti, 2013; Sussman & Sekuler, 2022) suggesting that technology accelerates the pace of life by increasing task switches and density, intensifying feelings of being rushed. This effect was especially noticeable for email and work communication apps, which likely heighten urgency and workload (Bittman et al., 2009; Mullan & Wajcman, 2019; Pritchard & Symon, 2023). Evidence for the latter intensification shows in the mediation through juggling load. This observation underscores the mechanism of mobile communication intensifying the (perceived) work load through the multitasking and switching between tasks, a finding consistent with literature on work fragmentation and role switching (Cornwell, 2013; Lu, 2024). Despite small effect sizes, these results are meaningful in the larger context of mobile communication's impact on daily life: they might accumulate over time, potentially leading to a more meaningful long-term impact on our (mental) well-being.

Social media use showed negative, non-significant associations with feeling rushed, possibly reflecting leisure usage. While a negative relationship aligns with the findings from Ross et al. (2023) — who found that social media use was often linked to *decelerated* time perceptions — further research with more data is needed to support this effect. While we did not find a direct link between social media use and feeling rushed, we did find a positive association between the duration, frequency and fragmentation of mobile social media use and juggling load. This finding calls into question whether there may be reversed causality, both here but similarly for the other categories, where in situations of increased juggling load, individuals turn more to social media (for relaxation) and mobile communication (for managing task loads). Future research is needed to unpack the temporal order of these associations.

A strength of this study lies in examining different features of mobile communication across various categories, revealing differential patterns. This supports the call by mobile media scholars to move beyond monolithic 'screen time' measures and consider the nuances of daily communicative behavior (de Segovia Vicente et al., 2024; Ross et al., 2023; Vaid et al., 2024). Our decision to classify app use by category was informed by previous work highlighting distinct connotations between these categories (e.g. Aalbers et al., 2023; Meier & Reinecke, 2021). However, we acknowledge that this classification carries inherent multidimensionality: app categories do not map cleanly onto communication modalities nor activities. For instance, the use of certain messaging applications may be used for both casual chatting and work-related coordination. Future research might explore other levels of CMC, such as message content, to more accurately distinguish different dimensions of mobile communication.

This study found associations at both within- and between-person levels. A third aim was to examine if heterogeneity was present in the size and direction of these associations, and could be explained by age, parenthood status and segmentation preference (as theoretically relevant explanatory factors of this variation). Individuals who engaged more with work communication apps and chat apps felt more rushed. This feeling was particularly pronounced among women, working individuals, and younger participants. This suggests stable differences, linked to social roles and occupations, indicating that mobile communication preferences and behavior can reflect social structures and display the broader 'habitus' of social groups (M. M. P. Vanden Abeele & Nguyen, 2024). Prior research also suggests that a harried lifestyle, potentially displayed through smartphone use, can act as a status marker (Fast et al., 2021). Future research may seek to investigate this further.

Although this study has offered new insights in the relation between mobile communication, feeling rushed and juggling between domains, it has limitations. First, the observed effect sizes of within-person effects were small, which may be related to the complexity of feeling rushed as a construct. This experience is likely influenced by multiple factors besides mobile communication, such as being at work which was a far greater predictor. Moreover, the experience of feeling rushed itself can be seen as containing multiple dimensions (Denovan et al., 2024).

Second, our operationalization of the juggling load may not fully have captured the complexity of juggling roles within a single domain. The similar levels of being rushed experienced when juggling within work versus across work and private domains suggests that the work role alone can be intense, with individuals often being required to switch between different sub-roles (Ashforth et al., 2000). It also suggests that our theory about cross-domain juggling constituting greater task load (relative to within-domain juggling) may not be sufficiently nuanced.

Third, this study did not examine other devices (e.g., laptops, desktops) commonly used in knowledge work, potentially affecting boundary management. Future research should explore how different devices impact perceptions of time pressure independently and how their combined use might even intensify these effects.

Fourth, our sample was limited to Android users, potentially influencing generalizability. Recruitment was conducted in collaboration with a large local newspaper and targeted an older, educated demographic. The voluntary nature of participation may also have introduced some bias. However, as there was no financial incentive, the data quality and reliability likely remain high.

Fifth, participants had the flexibility to engage with the study at their leisure, potentially limiting the experienced range of feeling rushed in daily life. We acknowledge that this selection bias can contribute to the low averages of the rushed measure. Indeed, a few participants explicitly communicated that they felt they lacked sufficient time to participate fully and chose to drop out. However, our analyses primarily focused on within-person effects, providing most insight into momentary, within-person dynamics, rather than generalizable, aggregate findings.

Sixth, the use of ESM might have, inadvertently, increased participant's feelings of being rushed or their juggling load, by adding additional task(s) throughout their day. In addition to this increased task burden, responding to repeated ESM prompts may have heightened participant's awareness of time — similar to how mobile devices themselves can function as constant reminders of the current time, through device's clock, timestamped messages, or calendar reminders (Ross et al., 2023). Together, this increased awareness of passing time may contribute to experience of being rushed, even in the absence of increased task density or task switching. Although ESM remains crucial for capturing momentary assessments, minimizing memory biases and maximizing ecological validity, future studies should account for its potential impact.

Finally, due to the observational design, reverse causality cannot be excluded (e.g., feeling rushed prompting greater chat use). However, such reversed effects seem less likely for externally driven features such as receiving email notifications. Future research should explicitly address these potential bidirectional relationships.

In conclusion, this study has contributed to the understanding of how different indicators of mobile communication — especially those related to work — are associated with feeling rushed, mediated by the experience of juggling multiple roles and tasks. Using a large sample of 774 participants, and by combining both passively logged smartphone data with ESM, we have attempted to capture both the dynamic and person-specific nature of these effects. Our findings support acceleration theory by empirically highlighting mobile devices' role in accelerating daily life. Lastly, we revealed that these associations can vary based on individual factors such as age, parenthood and segmentation preference, underscoring the importance of considering both situational and personal factors when investigating mobile communication's impact on wellbeing.

Computers Matter

The Relevance of Computer Log Data for Understanding Associations Between Digital Media Use and Well-Being[12]

[12]Manuscript under review as: Van Gaeveren, K., Murphy, S., de Segovia Vicente, D., & Vanden Abeele, M. (2026). *Computers matter: The relevance of computer log data for understanding associations between digital media use and well-being* [Manuscript submitted for publication]. Department of Communication Sciences, Ghent University.

Abstract

Research on digital media and well-being increasingly relies on smartphone log data to overcome limitations of self-reports, but often ignores computer use, implicitly treating it as less relevant. This study examines this assumption using experience sampling data from 106 participants who reported mood six times daily over two weeks (7,442 observations) alongside passive smartphone (499,696 events) and computer (257,362 events) logs. Results show that including computer use in analytical models significantly alters the pattern of findings. Including computer use increased individual and aggregate screen time estimates by an average of 136% relative to smartphone measures. Cross-device behaviors, such as sequential and overlapping use, were widespread, with participants switching devices on average 66 times per day. Models predicting daily and momentary mood were substantially altered when computer and cross-device metrics were included. These findings demonstrate that computers matter and underscore the need for multi-device data collection to improve validity of digital media effects research.

Keywords: digital trace data, computational methods, smartphone use, computer use, behavioral data, well-being

Introduction

Media scholars increasingly leverage smartphone log data to test theoretical claims about the impact of digital media use on health and well-being. This methodological shift is largely motivated by a desire to overcome the limitations of traditional self-report methods, such as recall bias, social desirability, and inconsistencies in self-reporting, which threaten the reliability and validity of findings (Sewall et al., 2020; M. Vanden Abeele et al., 2013). Indeed, a growing body of literature demonstrates that self-reports of smartphone use are inaccurate: When comparing self-report data with log data, for instance, studies find only moderate correlations concerning the duration and frequency of smartphone engagement (Molaib et al., 2025; Parry et al., 2021; Tkaczyk et al., 2024). This is partly due to smartphone use often taking place in the form of short and frequent interactions, and this across multiple apps or platforms, which makes detailed assessments of usage frequency and duration difficult to provide (Parry et al., 2021). Moreover, studies also show that the above inaccuracies in measurement are not random but vary systematically. For instance, heavy users are more likely to underestimate their use and vice versa (cf. regression to the mean (M. Vanden Abeele et al., 2013)), and self-report estimates are more inaccurate among individuals with lower psychosocial well-being (Sewall et al., 2020).

Because solely relying on self-reports of digital media use can lead to distorted findings and misleading conclusions (Kaye et al., 2020; Sewall et al., 2022), researchers increasingly turn towards passive monitoring of digital media behavior as the 'gold standard'. After all, the advent of device logging, combined with the computational turn in the social sciences, provides a unique opportunity to address several of the limitations of self-report data. Passively gathered device data offer richer, more reliable insights into media behaviors while also reducing participant burden. For instance, passive smartphone logging allows for precise measurement of smartphone usage patterns, both temporal and at the app-level, and this in a naturalistic setting (Ryding & Kuss, 2020). Collecting log data over extended periods of time also enables researchers to investigate within-person fluctuations, allowing for more nuanced and dynamic relationships to be modelled (Marciano et al., 2022; Sewall et al., 2022), thereby taking into account different temporal resolutions (Klingelhoefer et al., 2025). Moreover, researchers can identify and extract more complex and granular patterns from the data, such as features representing the fragmentation and 'stickiness' (Aalbers et al., 2021; Brinberg et al., 2023; Deng et al., 2019; Siebers et al., 2024) or 'burstiness' (Jo et al., 2012; T.-Q. Peng et al., 2020) of media use.

The above evolutions are important because, to date, there has been a tendency among researchers to rely on a limited set of cruder operationalizations when leveraging log data. More specifically, scholars often aggregate detailed log data into metrics such as daily screen time or frequency of use and incorporate only these aggregates as predictors in their analyses. Although useful for capturing broad trends, this approach risks obscuring important nuances in digital behavior (M. M. P. Vanden Abeele, 2021) and may increase the likelihood of type I (mistakenly detecting effects where none exist) and type II (not finding effects that *do* exist) errors. Alternative strategies, such as analyzing app-specific usage, exploring different temporal resolutions (e.g. fragmented vs. sustained use), or conducting sensitivity analyses, could provide a more robust test of the underlying theory. Rodman et al. (2024) provide empirical support for this claim, demonstrating that while overall smartphone use showed very weak or no associations with well-being, more granular features—such as engagement with distinct categories or platforms—were more predictive. Differentiating between distinct digital activities is seen as a crucial prerequisite for studies incorporating screen time measures. This is also argued by Meier & Reinecke (2021) who warn against treating digital media as a homogenous construct and failing to account for differences in applications, features and interactions. Similarly, Kaye et al.

(2020) propose shifting away from screen time towards screen *use*, focusing on the behaviors facilitated by digital devices, rather than simply measuring the time spent on them.

In sum, the increasing use of log data is generally considered a major leap forward for the field of research on digital media effects (Parry & Klingelhoefer, 2025). However, despite bringing obvious benefits, there is currently one major caveat that may compromise robust tests of theory, namely that log data are often collected from only one, or a limited type of devices, most commonly the smartphone, potentially leading to an incomplete picture of individuals' overall media behavior. Several studies have highlighted this limitation. For instance, a meta-analysis by Parry et al. (2021) found that nearly all studies investigating log-based digital media use relied exclusively on smartphone data, with only a handful of studies incorporating passive sensing data on the use of computers, gaming devices or general internet use. Liu et al. (2024) also highlighted the importance of distinguishing between the type of digital media use when assessing their impact on well-being but did not account for differences in device use. Similarly, Verbeij et al. (2022) found that digital log measures of social media were as predictive as self-reports but explicitly acknowledged that the log data were limited to smartphones, thereby excluding any usage on tablets, laptops or web-based versions of social media platforms.

This methodological limitation raises concerns: Research relying solely on smartphone data may underestimate people's digital engagement and may be blind to how particular (cross-)device features might potentially impact well-being. For instance, in another illustrative study, Schenkel et al. (2024) found that smartphone *email* app use was negatively associated with morning positive affect, while *chat* app use was negatively related to evening positive affect. But, as the authors also acknowledge themselves, by focusing on smartphone data alone, the study likely overlooks important contributions from other devices, most notably the computer or laptop, which especially in an adult, employed sample are devices crucial for many users in managing email and chat communication. Indeed, daily computer use is extensive - in many instances likely comparable to or even surpassing smartphone use. Thus, to regard not capturing computer log data as a rather minor limitation seems problematic. By failing to capture computer-based log data alongside smartphone log data, the results and conclusions of research may not just be a *little* inaccurate—such that point estimates are a little off, or the confidence intervals a little wide—but rather *very* inaccurate, such that point estimates can be *way* off, and even in the opposite direction. Moreover, there seems to be a dearth of research highlighting the *affordances* a multidevice approach may offer, as there is 'media multiplexity' in people's daily media repertoires (Haythornthwaite, 2005). By additionally capturing computer screen time, for instance, it becomes possible to identify device switching dynamics and tendencies to employ multidevice workflows. These insights could have various theoretical insights, for instance in the context of novel theorizing around media multitasking (Abrahamse et al., 2025).

Accordingly, the aim of the present study is to more clearly underline the true extent of the impact, both on research findings and theory, when computer log data is leveraged alongside smartphone log data. We do this by first describing how accounting for computer/laptop use log data changes 'screen time' measures. Second, we examine whether—and how—the inclusion of these computer-based log data (both as stand-alone predictors and in combination with smart-phone log data (i.e. cross-device features)) might affect conclusions regarding the relationship between digital media use and individual outcomes. To prove our point, we leverage data from a large-scale, two-week experience sampling study, that measured subjective wellbeing (and other) constructs six times a day while logging participants smartphones and computer use. In this study, we specifically examine the mood data (i.e., how happy and 'down' participants felt) that was captured, given several studies have explored associations between passively sensed smartphone use and mood (Balliu et al., 2024; Pratap et al., 2019), but without accounting for computer/laptop use.

Methods

Participants

This study draws from an experience sampling project that recruited 1,315 participants to investigate the link between digital media use and wellbeing, which was conducted from October to December 2022. Here, we leverage data from 106 participants (Male: 49%, Female: 48%, other: 3%), with a mean age of 39.4 (SD = 11.0), most of whom were highly educated (Bachelor's degree or higher: 94%, High school degree: 5%, Lower: 1%) and in full-time employment (75%, Student: 7%, Other: 18%), that *additionally* provided smartphone and computer log data alongside their ESM data. These participants contributed 499,696 smartphone, 257,362 computer, and 7,442 ESM datapoints.

Procedure

The project received approval from the IRB of Ghent University Faculty of Political and Social Sciences (code 2022/11). Written (digital) informed consent was asked during the onboarding process: when signing up, and during installation of each smartphone app used in this research. Consent was required before any of the research tools could send out questionnaires or gather data. All study procedures were conducted in accordance when Ghent University guidelines.

Participants registered for the research online, and installed two smartphone applications[13]: one that provided them with ESM questionnaires (m-Path; www.m-path.io), and one that monitored their smartphone log data (MobileDNAPlus; e.g., screen time, number of notifications received). Participants also installed ActivityWatch (https://activitywatch.net/) on their computer, which passively captured their screen time.

During their two-week participation, six questionnaires were sent out daily via the m-Path app, between 07:30-09:00 hrs, 10:15-11:45 hrs, 13:00-14:30 hrs, 15:45-17:15 hrs, 18:30-20:00 hrs, and 21:15-22:45 hrs. These time slots were separated by a period of at least 1 hr 15 min and at most 4 hr 15 min. Following the initial notification, each ESM questionnaire remained available to complete for 45 min. Participants received a reminder notification if the questionnaire had not been responded to 30 min after the initial notification. Participants answered the same mood items during each ESM questionnaire they completed. Throughout the two-week period, the MobileDNAPlus app and ActivityWatch tracked screen time events: timestamped app use and notifications on the smartphone, and timestamped computer program use on their computer. Log data were processed into different features and aggregated at two time-intervals: at the ESM questionnaire window interval, and at a daily interval.

Measures

Self-report

Mood was measured using two items, both containing the stem "Since getting up this morning…" (first questionnaire received each day) or "Since the last beep…" (the five remaining questionnaires received each day): "…I felt happy" and "…I felt down.". Both items had a 7-point Likert scale ranging from 1 ('Not at all') to 7 ('All the time').

Smartphone and Computer Log data

[13]Detailed methodological information provided on https://osf.io/fzmjx/overview?view_only=3e80e078d4354722a2a6773cb6479d34. Here, only pertinent methodological information provided.

Device usage was aggregated into separate measures per device, and per category of use. All smartphone apps, computer programs and browser titles were categorized, of which we included four labels in this study: email (e.g.: Outlook, Gmail), work (e.g.: Excel, Word), social (e.g.: Facebook, Instagram), and chat (e.g.: WhatsApp, Signal).

Duration sums the total screen time, and total screen time per category. **Frequency** sums the total amount of app launches, and again per category. **Fragmentation** looks at the dispersion of these two dimensions across time: A low value signifies less time was spent on the device/ category and this across fewer usage sessions, while a high value indicates that more time was spent and this in frequent and short bursts (refer to Siebers et al., 2024 for more information). In total, this led to forty-five different variables: three features across five categories (total, email, work, social, chat) and three device (smartphone, computer, combined) levels.

Additionally, we calculated three cross-device features, based on both smartphone and computer log data. Figure 5.1 contains an example of how each feature was calculated.

Device switching, which counts the number of times a participant alternated between devices during a combined usage session, or when they stopped a device session and immediately started a new usage session on a different device. Our example shows three device switches taking place, twice originating from smartphone use (1, 3) towards computer, once from computer use (2) towards smartphone.

Context switching, where we counted the number of times a device switch also meant a change of context, i.e.: the category of use on device A was different than the category of use on device B. For example, while having an email client active on the computer, the number of times a non-email smartphone session was started. Our example shows a longer period of email use on the computer, and three separate smartphone sessions, of which two were seen as context switching. The first from email to social media, the second from email to chat.

Overlapping device use, which takes the proportion of device time where device A and device B were active at the same time. For example, take a one-hour time window which a user spent working on their computer, and during this time spent six minutes using social media on their smartphone, the computer overlap percentage would be 10%, while smartphone overlap percentage would be 100%. Finally, the figure shows the same example of device switching, but this time highlights the period where we consider overlapping device use to be present. Both computer and smartphone use total eight minutes, of which one minute overlaps. Hence, 12.5% of screen time would here be considered as overlapping.

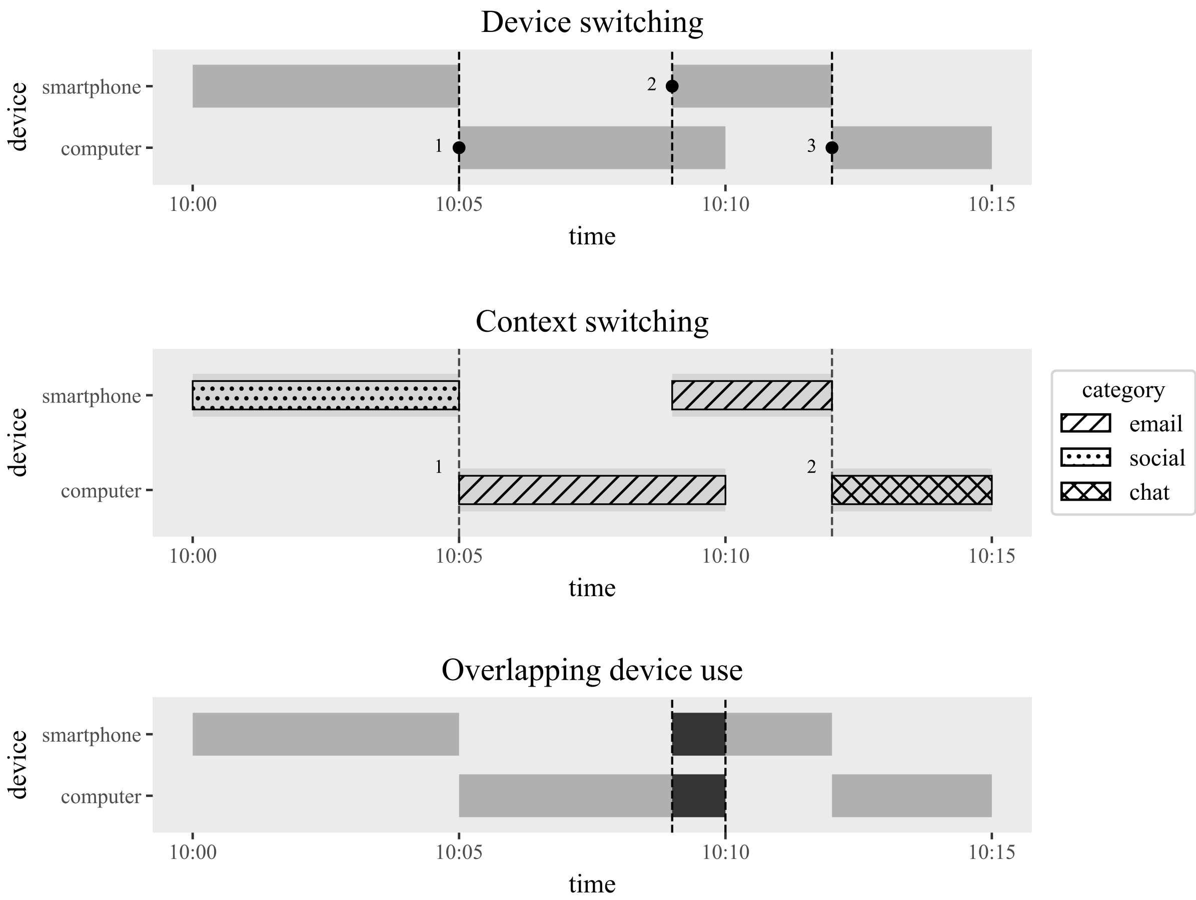


Figure 5.1: Examples of each cross-device feature.

The top panel illustrates device switching, showing sequential transitions between smartphone and computer use. The middle panel illustrates context switching, where a device switch coincides with a change in category of use (e.g., from computer email use to smartphone chat or social use). The bottom panel illustrates overlapping device use, showing periods where both devices are active simultaneously.

Data Analysis

Data processing and preparation was carried out using the Python Polars package (The Polars Development Team, 2025) and dplyr R package (Wickham et al., 2025). Descriptive statistics were calculated using the same Polars package. This resulted in a momentary dataset – where each feature was calculated per ESM time window, which were semi-randomized with an average of 3h 4m in between – and a daily dataset, where each feature was calculated at a daily time window.

Associations between Mood and Device Use were modeled using Linear mixed-effects models in the R package lme4 (Bates et al., 2015). Momentary models used the raw mood variable (feeling happy, feeling down), while device use variables where within-person scaled (i.e.: each participant's average was 0, with a standard deviation of 1) to allow for easy comparison. Daily models used the average daily mood and again within-person scaled device use variables.

We employed two tailed testing and considered effects significant if p-values were .05 or less.

Transparency and Openness

Materials and code used for this study can be accessed on our Online Science Framework (OSF) page: https://osf.io/fzmjx/overview?view_only=3e80e078d4354722a2a6773cb6479d34.

Results

In this section, we demonstrate how including computer-based log data reshapes both the descriptive conclusions about digital media use and inferential conclusions about its associations with mood. We start of by showing how total and categorical estimates of screen time change markedly when computer data are incorporated. Next, we show how employing a multi-device perspective reveals insightful cross-device dynamics (from device switching or multitasking), which cannot be gleaned from smartphone data alone. Finally, we demonstrate how additionally including computer log-data impacts inferences about the relationship between media use and affect.

Screen Time is More Than Smartphone Time

To date, most research incorporating device log data has predominantly operationalized 'screen time' as the aggregate of time spent on the smartphone. Our findings reveal that this assumption underestimates people's digital engagement and misrepresents how screen time is distributed across devices. Within our sample, for instance, participants spent an average of 2h 5min interacting with their smartphone (see Figure 5.2), with the minimum and maximum daily average being 13min and 6h 22min, respectively. However, within our sample, participants spent an average of 2h 10min interacting with their computer, an amount essentially on par with smartphone use. Considering workdays only, participants used computers more still, with screen time averaging 2h 20min per day, while smartphone screen time dropped to an average of 1h 51m. For daily computer use, minimum and maximum were 2 minutes, and 12 hours and 13 minutes daily.

But most important is the impact of capturing *both* smartphone and computer screen time. Our analyses reveal that participants spent an average of 4h 15min interacting with both their smartphone and computer combined, which represents an 136% increase in screen time relative to the smartphone alone. Interestingly, our analyses also reveal more variance in participant screen time relative to logging the smartphone alone (SD = 136.1 vs 68.5). Moreover, our analyses reveal that just because participants have similar smartphone screen time, that it does not follow that their overall screen time will be similar. Indeed, Figure 5.2 shows four participants from our sample with near identical smartphone screen time, displaying very different computer use time, and thus combined screen time: For some participants, such as the participant highlighted in red, including a second device marks a large increase in total daily screen time, while for others – such as the participant highlighted in yellow – these changes are less pronounced.

In sum, our findings concerning screen time alone represent a critical reason for additionally logging participants' computers. They not only show that sample and population screen time estimates will be substantially underestimated without computer logging, but they also demonstrate that participant patterns within devices may paint a markedly different picture. On this point, if screen time *did* exert a negative effect on wellbeing, a participant that gets most or all of this via their smartphone would contribute valuable data to a study that only logged smartphones. In contrast, a participant that gets much of their screen time through their computer would add a lot of error, and more problematically, bias to the model, for a substantial amount of their screen time simply would not get measured. However, differences in screen time represent but a small part of this story.

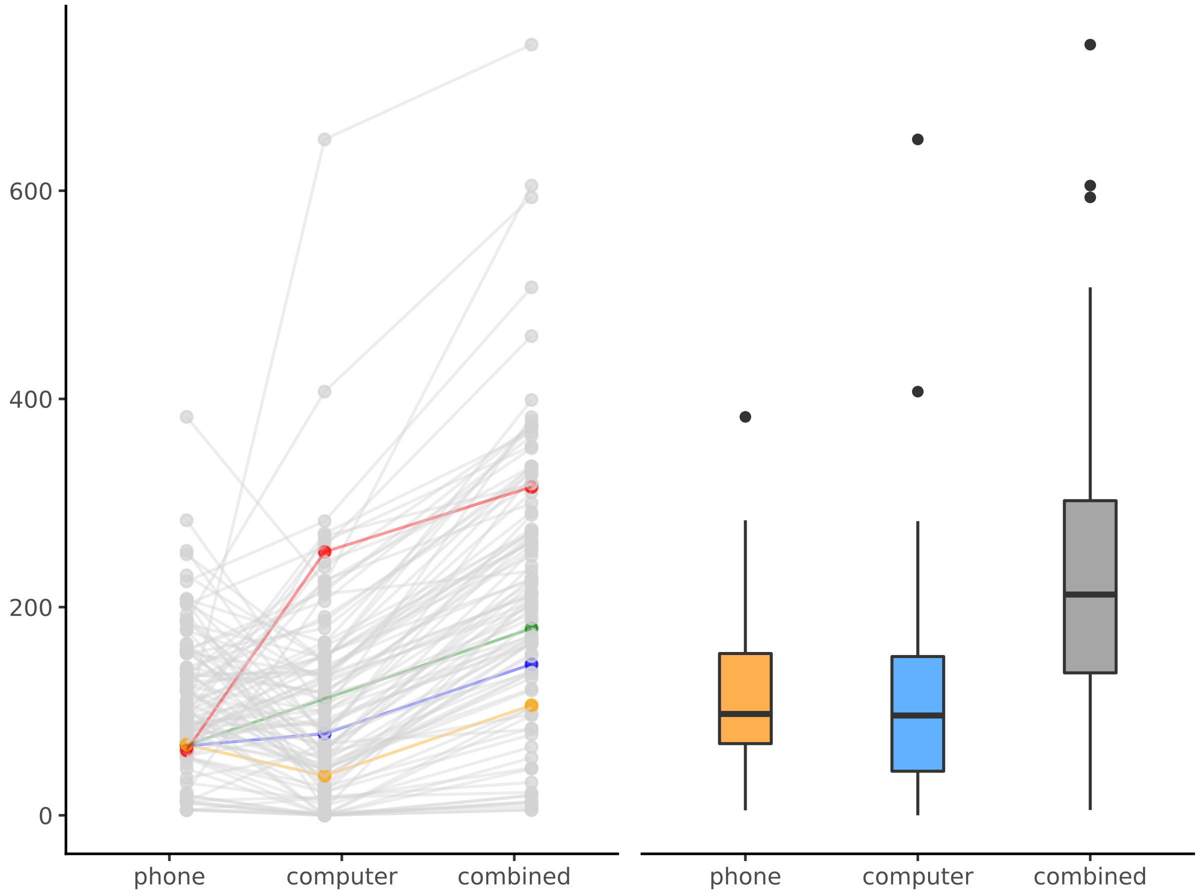


Figure 5.2: Average daily screen time per device, with left four participants highlighted.

The left panel shows individual-level averages for each participant, with lines connecting each participant's smartphone, computer and combined screen time, visualizing within- and between-person differences. Four participants with similar smartphone screen time are highlighted in color to illustrate how overall screen time diverges once computer use is included. The right panel displays boxplots summarizing the distributions of participant screen time across devices, highlighting differences in central tendency and variability.

Devices Shape What We Measure

Including computer log data also markedly impacts the story of *how* people engage with specific categories or platforms. For instance, imagine we are interested in how spending time emailing impacts our mood. It is possible, given that smartphone logging is so easy relative to computer-based logging, to believe one may be able to gain valuable insights from smartphone log data alone. As Figure 5.3 highlights, this may be true for participants who deal with their emails primarily using their phone. However, for the majority that use their computer to receive and deal with a sizeable number of their emails, any analyses would be severely hindered from the outset. The same goes for research that attempts to gain insights on the impact of emails exclusively from computer-based logging, given the increased ease at which emails can be accessed and dealt with via the smartphone. Overall, our analyses reveal that if we relied only on smartphone log data, we would be missing out on more than half of all email screen time (mean phone = 42.2%, mean pc = 57.8%).

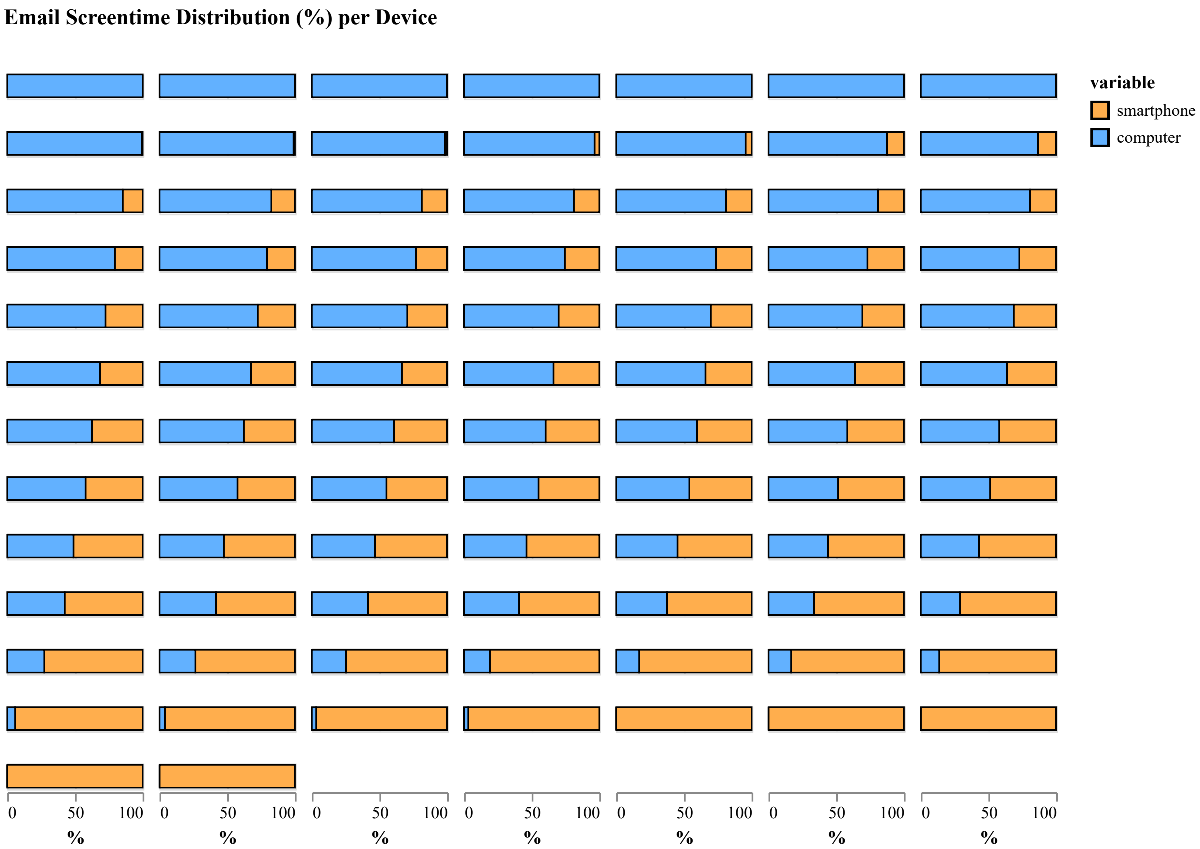


Figure 5.3: Email Screen time Distribution (%) per Device.

Each horizontal bar represents a single participant, with blue segments indicating the proportion of computer email screen time and yellow indicating the proportion of smartphone email screen time. Participants are ordered by their proportion of computer vs smartphone screen time. The figure demonstrates substantial between-person heterogeneity, with a large proportion of total email screen time taking place on the computer. This includes the 86 users of whom we had at least 3 days of smartphone and laptop data (each horizontal bar represents a single individual's distribution).

Taking additional app categories into account, we find large differences in the proportion of their screen time over the two examined devices: social media (mean phone = 62.9%, mean pc = 37.1%), chat (mean phone = 92.1%, mean pc = 7.9%), and work (mean phone = 24.9%, mean pc = 75.1%). This picture is similar when investigating specific platforms like Facebook, X, or Instagram. As an example, Figure 5.4 shows a timeline plot of a participant's engagement with Facebook during a single day. We see that, although they begin using Facebook on their computer, they frequently switch towards their smartphone in the afternoon and evening. Including additional devices allows us to investigate these different media habits more comprehensively. From a theoretical standpoint, this matters because when we operationalize 'social media use' or 'email use' solely through smartphone data, we could ignore substantial and distinct proportions of those media use practices.

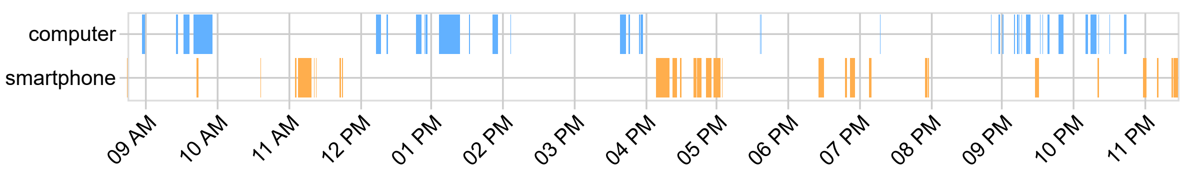


Figure 5.4: Timeline of Facebook usage across devices.

Vertical bars indicate periods of active use on the computer (blue) and smartphone (yellow), illustrating how social media engagement is not limited to a single device and highlighting the cross-device nature of this type of activity.

Beyond Duration: Cross-Device Dynamics

Apart from more accurately assessing someone's screen time and app usage, including multiple devices allows us to show how people juggle these devices throughout the day, coordinating tasks and attention across screens. Specifically, we can investigate patterns of sequential and multi-tasking use, assessing how often people switch devices and contexts, what proportion of their device's screen time overlaps, and how often they engage in these types of behaviors. Doing so may offer important theoretical insights that screen time alone is unable to provide, and is in line with recent calls to go beyond aggregate measures to ensure a better test of theory (Parry & Klingelhoefer, 2025). For more information on how we calculated these metrics, please refer to the Methods section.

First, we examined how often participants transition from one screen to another. On average, participants switched devices approximately 66.15 times per day (SD = 54.80, median = 52), with this behavior occurring on around 60% of all study days. This frequent toggling underscores that media engagement is rarely limited to a single device. From a measurement standpoint, this means that studies focusing on a single device likely capture only fragments of a broader digital routine.

Next, we examined overlapping use, or moments where both devices were active at the same time. Despite occurring less frequently than sequential use, overlap was far from rare: on days where participants used both devices, 18.4% (SD = 24.1) of computer use and 14.2% (SD = 17.2) of smartphone use was concurrent with the other device. This finding suggests that multi-device multitasking, such as working on the computer while simultaneously checking your smartphone, is a common behavioral pattern. This overlap reveals not just concurrent attention, but also potential concurrent cognitive or affective demands, which may influence well-being in ways single-device data cannot detect.

We further analyzed context switches, capturing moments where a device switch also entailed switching between categories of media activity (e.g.: from computer email software to a smartphone chat application). On average, participants made 59.45 context switches per day (SD = 49.87, median = 46), indicating frequent alternation between platforms and activities. Notably, nearly 90% of device switches also involved a context switch, meaning that device transitions almost always implied shifting what one was doing, not merely where. This high coupling of device and context changes highlights that device switching can often entail a reorientation on cognitive or situational levels, potentially shaping attentional and emotional outcomes.

Taken together, these cross-device metrics reveal that digital media use is not only distributed across devices but dynamically structured in time. By capturing switching, overlap and context transitions, we can uncover a layer of media behavior that single-device studies cannot observe. This not only improves our understanding of how people engage with screens, but also enables more precise tests of theoretical mechanisms, such as multitasking or attentional fragmentations, that depend on the interplay between devices.

Having established these dynamics, we next examine how they relate to affect well-being, testing whether these multi-device patterns help explain momentary and daily mood fluctuations.

Multiple Devices, Multiple Stories

Having established that computer use is both substantial and intertwined with smartphone activity, we now examine how these patterns translate into differences in affective experience. Specifically, we test whether including computer-based and cross-device features change the observed relationships between digital media use and mood. In doing so, we show that the scope of data, whether limited to smartphones or extended to multiple devices, can influence both the direction and strength of media effects estimates. Figure 5.5 and Figure 5.6 visualize how the inclusion of computer and combined device data reshape associations with mood.

Models using only smartphone data revealed few and inconsistent associations between digital use and mood. For happiness, only a handful of momentary associations emerged, such as minor negative links with email and work-related use, and a small positive link with chat frequency. In addition, all day-level associations were non-significant. Similarly, models predicting feeling down found some positive associations with total, email and social smartphone use, but these effects were small in magnitude. At the daily level, small effects for total, email and work-related use were also found. These patterns replicate what we already know from prior research using smartphone log data: There is limited evidence for strong mood effects, and effects are inconsistent across categories of use.

When we adapted our models to use the computer-based metrics, the picture changes notably. At the momentary level, all computer features showed a significant negative association with feeling happy, while at the day-level these associations remained for total, email and work categories. When modelled for feeling down, significant positive associations were found for all categories of use, except chat, both at the momentary and day-level. In other words, on days and moments of heavier computer use, participants reported slightly lower positive affect. This pattern was more consistent and robust than those found in the smartphone-only models.

This divergence demonstrates that different devices capture different experiences and contexts. What appears neutral or weakly related to mood on smartphone may signal strain, workload or mental fatigue when observed via computers.

When smartphone and computer measures were combined, several effects changed significance. Daily total device frequency, which was negatively related to happiness in computer models was no longer significant in the combined device model. At the momentary level, combined social media use duration and frequency were no longer significant compared to computer models, while on the day-level fragmented patterns of combined social media and chat use now emerged as new, significant predictors of lower happiness. The fragmented nature of combined daily chat use now also showed a significant positive association with feeling down. This highlights the importance of looking beyond *how much*, and additionally consider *how fragmented* and *cross-device* screen time occurred. These shifts between device models underline that our conclusions about digital well-being are not fixed properties of the phenomenon but a result of how comprehensively we choose to measure it.

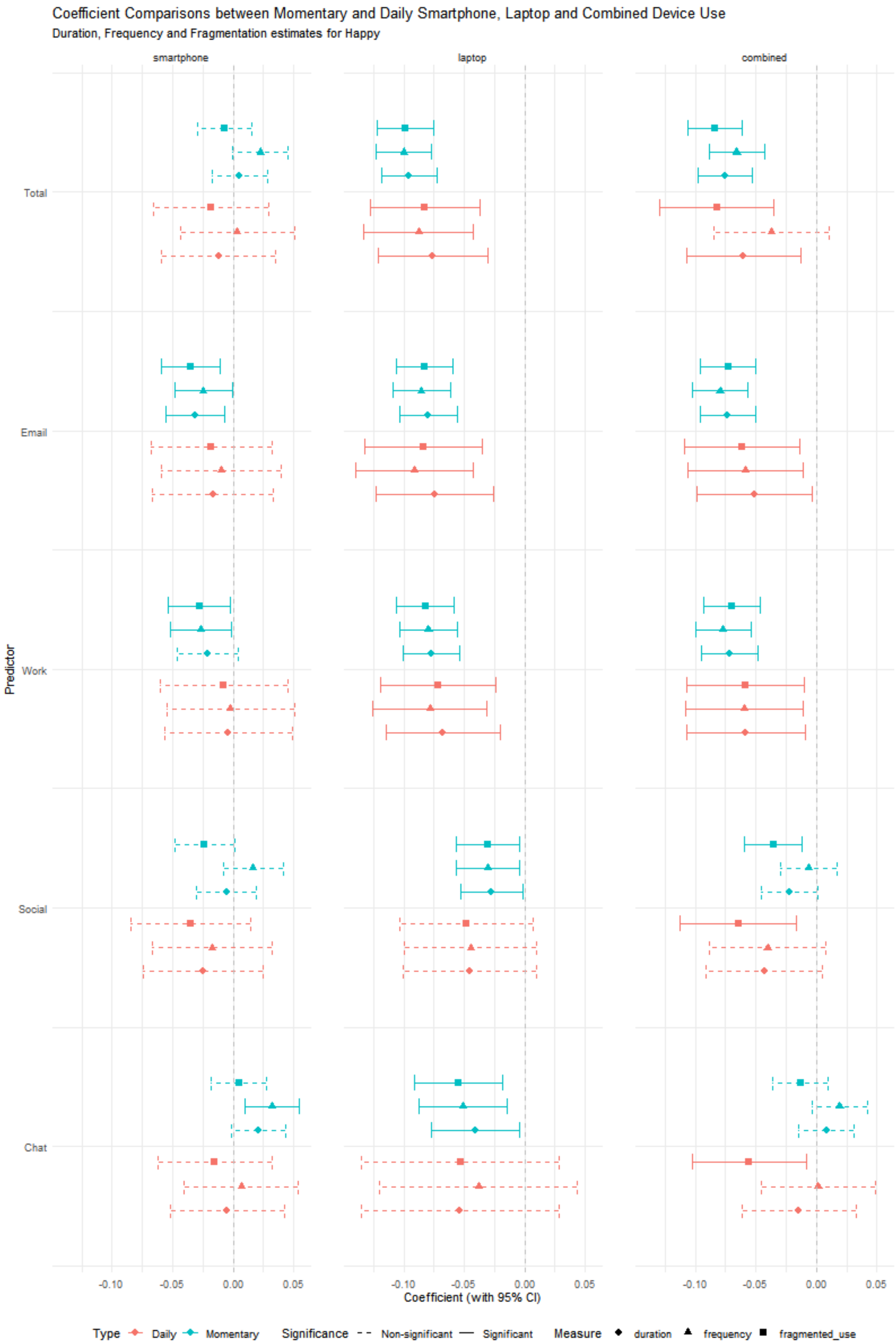


Figure 5.5: Overview of device use metrics' association with feeling happy.

The figure is organized in three columns by device type (smartphone, laptop, combined) and five rows by category of use (total, email, work, social, chat). Within each cell, standardized coefficients are shown for duration, frequency and fragmentation (diamond, triangle, square) at both the momentary (blue) and daily (red) level. Solid lines indicate significant associations; dashed lines indicate non-significant effects.

Figure 5.6: Overview of device use metrics' association with feeling down.

Mirrors Figure 5.5, reporting associations with feeling down across device types, usage categories, metrics and temporal levels.

Finally, we examined cross-device dynamics, which describe how devices are used together rather than separately. Table 5.1 and Table 5.2 contain the output of our cross-device models predicting feeling happy and down. Here, patterns were very consistent: more frequent device and context switches, as well as a greater overlap between smartphone and computer use were all associated with lower happiness and higher feelings of being down.

Table 5.1: Cross-device features and Feeling Happy.

Feature	β	95% CI	p	level
Switches	−0.070	[−0.093, −0.046]	.000	Momentary
Context switches	−0.068	[−0.091, −0.044]	.000	
Overlap % (phone)	−0.070	[−0.094, −0.046]	.000	
Overlap % (computer)	−0.039	[−0.063, −0.016]	.001	
Switches	−0.073	[−0.119, −0.027]	.002	Daily
Context switches	−0.079	[−0.125, −0.033]	.001	
Overlap % (phone)	−0.071	[−0.117, −0.025]	.003	
Overlap % (computer)	−0.027	[−0.073, 0.019]	.257	

Fixed-effect estimates from mixed-effects models examining the associations between cross-device use and happiness *at the momentary and daily level. Standardized coefficients, 95% confidence intervals and p-values are shown.*

Table 5.2: Cross-device features and Feeling Down.

Feature	β	95% CI	p	level
Switches	0.035	[0.012, 0.057]	.003	Momentary
Context switches	0.031	[0.008, 0.053]	.007	
Overlap % (phone)	0.024	[0.001, 0.046]	.037	
Overlap % (computer)	0.029	[0.007, 0.052]	.011	
Switches	0.069	[0.031, 0.107]	.000	Daily
Context switches	0.079	[0.041, 0.117]	.000	
Overlap % (phone)	0.049	[0.010, 0.087]	.012	
Overlap % (computer)	0.047	[0.009, 0.085]	.016	

Fixed-effect estimates from mixed-effects models examining the associations between cross-device use and feeling down *at the momentary and daily level. Standardized coefficients, 95% confidence intervals and p-values are shown.*

These results suggest that the coordination of digital activity across devices, through rapid transitions, multitasking and overlapping attention, may be more psychologically consequential than total use duration. The negative associations align with theories linking attentional fragmentation to cognitive depletion and possible reduced emotional recovery (Brinberg et al., 2023; Siebers et al., 2024). Importantly, such patterns could only be observed through multi-device data.

In sum, including computer and cross-device data altered the inferential conclusions we could draw. This demonstrates that the scope of measurement, the devices and dynamics included, shape the conclusions we draw about media use and wellbeing. By moving from single-device to multi-device models, we obtain a fuller picture of digital behavior, as well as a different understanding of how this behavior related to emotional outcomes.

Discussion

The main goal of this paper was to examine how the inclusion of computer log data alters conclusions about in digital media effects research. Research on media effects has increasingly embraced smartphone-based log data to overcome the limitations of self-reports, yet this smartphone-centric approach risks underrepresenting people's total digital engagement and obscure how people's media behaviors take place across devices. By integrating smartphone ánd computer log data, and combining these data with experience-sampled mood reports, we highlighted that extending the measurement scope to multiple devices, both descriptive estimates of screen time as well as inferential patterns linking media use to affect can drastically alter. These findings demonstrate that the conclusions we draw about emotional consequences of digital media use will depend on which parts of a persons' device ecology we capture.

With our descriptive findings, we have shown that computers matter: The amount of screen time spent on these devices, especially in the case of knowledge workers, greatly contributes to the total engagement with digital devices, even eclipsing the time spent on the field's main focus, the smartphone. In our sample, total screen time roughly doubled when computers were included. Beyond this overall increase, including a secondary device revealed notable domain-specific differences. Certain categories of use – such as email or work – were much more present on the computer. Even in categories of use typically related to smartphones we saw large cross-device activity. For example, more than one third of all social media screen time occurred on computers, illustrating that people often distribute the same activities across multiple devices, rather than confining it to a single one.

The inclusion of computers does more than change how much - or what kind of - media use we measure; it allows us to reshape our understanding of how people use these devices. As research on, for instance, media multiplexity (Haythornthwaite, 2005), engagement with multiple devices is not neatly partitioned in separate sections, but rather dynamically shifting throughout the day. By introducing cross-device dynamics such as device switching, context switching and overlapping use, we can move beyond more "static" indicators of screen time and try and capture these more temporal dynamics. Our analyses revealed that cross-device use was not some occasional occurrence, but rather a defining feature of how people shape their daily digital practices. On average, participants switched more than sixty times per day, with these types of transitions happening on the majority of study days. Most device switches also entailed a context switch, meaning users changed both device and activity (e.g.: from composing an email on their computer to writing a chat message on the phone). Finally, overlapping use was also common. When both devices were active on a given day, nearly one fifth of computer use and one seventh of smartphone use overlapped with the other device. Together, these findings reveal a more dynamic pattern of digital behavior that single-device studies cannot capture, and is likely to shape how media use relates to affective well-being.

Our inferential results, finally, show that expanding the measurement scope from one to multiple devices meaningfully reshapes the conclusions we can draw about the link between media use and affect. In line with previous research, models using only smartphone data revealed few and inconsistent within-person associations with mood (Winbush et al., 2025). When computer-based metrics were added, however, much more consistent results were found for the adults in our sample: higher levels of computer use, and especially for work and email related activities, were associated with lower happiness and higher feelings of being down. Importantly, cross-device features showed similar and even slightly stronger associations with mood, indicating that coordinating activities across screens may be more consequential that total use alone.

Of course, our study is not without its limitations. First, participation in computer logging was an optional step in the broader data collection process, introducing an element of self-selection bias. It is very likely that participants who opted to install the additional software may differ systematically from those who did, perhaps because they felt rely more on computers for work or perceive their computer use as more impactful for their digital well-being. This means our sample might overrepresent participants with strong computer engagement. Nevertheless, there was substantial heterogeneity within this group of participants: some provided data of their main work computers, others from private laptops or even gaming desktops, resulting in diverse and informative usage patterns.

Second, our analytic sample was small and consisted mostly of adult knowledge workers. We cannot claim these data are very representative of a broader population, even though the data were intensive in both time and measurement resolution. As the goal of this study was primarily methodological, illustrating how including additional device could alter the conclusions we draw from these types of data, rather than make population-level claims about specific effect sizes.

Third, our approach focused on the inclusion of a single, but crucial, device. While adding computers represents an important step towards a fuller account of people's total digital lives, it still leaves out tablets, wearables and other digital devices. As personal device ecosystems continue to expand, the question is not only whether but how far researchers should extend the boundaries of digital log collection. Addressing this complexity may require a shift towards far more individualized, within-person approaches. An n = 1 modelling perspective, where each participant's digital ecosystem can be modelled in relation to their own mood dynamics, could better take into account the heterogeneity of device configurations and media habits. Such idiographic models would allow researchers to detect patterns which hold for certain individuals and to identify which forms of digital engagement are most relevant for each user. This aligns with a growing call for more personalized and context-specific models of digital well-being (Valkenburg et al., 2021; M. M. P. Vanden Abeele, 2021).

Conclusion

In sum, our findings demonstrate that computers matter. Extending log-based measurement beyond the smartphone fundamentally changes both what we capture and what we conclude about media effects and mood. Multi-device data reveals that people's digital engagement are dynamic and distributed across screens, properties not adequately addressed in single-device approaches. Accounting for cross-device features improves not only the validity of media effects research, but also opens up new possibilities to model individual differences and bring in more contextual nuance in how our relationship with digital devices relates to our well-being. Embracing multi-device and person-centered approaches will be essential for building a more accurate understanding of digital media effects.

Data Availability

Anonymized data, materials and code used for this study can be accessed on our Online Science Framework (OSF) page: https://osf.io/fzmjx/overview?view_only=3e80e078d4354722a2a6773cb6479d34.

Author Contributions Statement

K.V.: Conceptualization; Data Curation; Formal analysis; Investigation; Methodology; Project administration; Resources; Software; Visualization; Writing - original draft; Writing - review and editing

S.M.: Investigation; Methodology; Project administration; Supervision; Writing - review & editing

D.D.: Investigation; Project administration; Resources

M.V.: Conceptualization; Funding acquisition; Investigation; Project administration; Supervision; Writing - review & editing

Funding

This study was funded by the European Union's Horizon 2020 research and innovation programme under the European Research Council (ERC) Starting Grant agreement "DISCONNECT" No. 950635, and by the FWO under COSTOF24/7 No. G075524N.

Making the Impossible Possible

Leveraging built-in features for non-intrusive and accurate Apple Screen Time tracking through ASTER[14]

[14]Published as: Martens, M., Van Gaeveren, K. *Making the impossible possible: Leveraging built-in features for non-intrusive and accurate Apple Screen Time tracking through ASTER*. Behav Res 58, 179 (2026). https://doi.org/10.3758/s13428-026-03065-2

Abstract

This paper describes a novel procedure to collect Screen Time data from iOS, iPadOS, WatchOS and MacOS devices using their built-in Apple Screen Time features. Traditional methods of studying digital behavior often rely on self-reported data, which are prone to inaccuracies (i.e., recall bias), limiting their validity and granularity. While Android devices have long facilitated real-time and granular behavior tracking through third-party applications, similar tools are not available on Apple devices due to Apple's restrictions. This is problematic because there are significant differences between the user populations of iOS and Android. To address this gap, this study developed a data donation procedure that leverages the synchronization of screentime enabling the extraction of comprehensive usage data of iPhones, iPads, Apple Watches and Macs linked to the same Apple ID. The process involves donating system-level files used to generate Screen Time metrics on Mac, containing anonymized use data of all linked devices. We developed a tool that enables researchers to process these files into a usable dataset (e.g., csv, JSON). This dataset provides granular insights into app usage without requiring substantial technical expertise or financial investment. While this approach enables the integration of Apple cross-device behavior into digital media research, it is limited to users with a Mac and can only capture data from the previous four weeks. Additionally, the method is vulnerable to changes in Apple's software structure echoing the moving target problem. Nonetheless, this method marks an important step forward in current approaches to the passive sensing of smartphone behavior.

Keywords: Passive Sensing, Apple, Screen Time, Monitoring, iOS, iPadOS, MacOS, WatchOS, digital phenotyping

Introduction

Measuring user behavior on mobile devices such as smartphones and laptops is crucial for academic research in the social and behavioral sciences, where a growing number of researchers attempt to unravel the correlates and effects of using digital media. For instance, how digital media use is linked to political participation and polarization (Chan & Yi, 2024), physical health (Nisafani et al., 2020), educational performance (Limniou, 2021), or mental well-being (Meier & Krause, 2023). To date, however, methodological challenges remain in the study of digital media use and its effects (Parry et al., 2022). A first challenge is that most research still relies primarily on cross-sectional self-report data. This traditional approach is often inadequate for capturing user behavior, as it lacks both specificity and validity (Araujo et al., 2017; Ohme et al., 2021; Parry et al., 2021):

With respect to specificity, self-reported measurements of media behavior, such as survey questions assessing daily smartphone use, afford only a coarse capture of online behavior; they lack the granularity needed to find clear relationships between these behaviors and outcomes of interest. Researchers may, for instance, be specifically interested in the fragmented nature of digital media use (Araujo et al., 2017; Griffioen et al., 2020), and/or may need access to the real-time dynamics of the behavior to understand its temporal associations with outcomes, such as time perception (Ross et al., 2023) or distraction (Siebers et al., 2024). This granularity becomes especially relevant when considering the use of passively sensed online behavior for the digital phenotyping of users' health and well-being (e.g., Aalbers et al., 2023), or for other forms of real-time predictive analytics that use digital trace data (e.g., for digital forensics).

With respect to validity, there is a known discrepancy between actual and self-reported (i.e. perceived) digital media behavior. A main reason for this discrepancy is recall bias (Parry et al., 2021; Scharkow, 2016; Sewall et al., 2020). Studies show that this discrepancy is systematic; it is greater, for instance, among individuals with psychosocial vulnerabilities (Sewall et al., 2020), and heavier users are more likely to underestimate the frequency of their behavior and vice versa (M. Vanden Abeele et al., 2013). There are also problems of social desirability in self-report responses (Ellis et al., 2019; Parry et al., 2022), and there is a risk of common method bias (Parry et al., 2022).

There is growing consensus among social and behavioral scientists that the above shortcomings in the accurate and valid measurement of digital media behavior severely limit the extent to which we can rely on self-report data for understanding how digital behavior affects individuals. Scholars are, therefore, increasingly responding to this issue by seeking alternative ways to measure digital media behavior, most notably through the collection of passively monitored digital behavior, which, albeit not entirely free from biases, is currently considered the 'gold standard' (Meier & Krause, 2023; Parry et al., 2022).

In the past five years, several studies have successfully measured objective smartphone behavior using both off-the-shelf and custom-built mobile applications (e.g., de Segovia Vicente et al., 2024; Johannes, Meier, et al., 2021). However, most available passive sensing apps are limited to Android devices. While Android-based sensing is feasible, it often requires substantial effort, expertise, and time to adapt processing scripts for specific research purposes. Other frameworks, such as Screenomics, rely on screen capture analysis (Reeves et al., 2021). Yet, none of these logging approaches work on iOS. This limitation stems from Apple's strict App Store policies, which prohibit apps designed to measure smartphone behavior, and from the Screen Time API's lack of specificity and sandboxed nature. The omission of iOS data in the research field is problematic (Baumgartner et al., 2023) given that more than 30% of smartphones run iOS

(De Marez et al., 2022). Moreover, Android and iOS users differ in sociodemographic profiles and psychological traits, with iOS devices generally positioned as premium products (Shaw et al., 2016).

Some scholars currently circumvent this problem using data donation procedures in which people donate screen captures of the built-in 'Apple Screen Time' features of the iPhone operating system (e.g., Baumgartner et al., 2023; Ohme et al., 2021). These screen captures can then be automatically 'read' using a custom-developed script that enters the information of the photo or video into a database. These solutions can collect aggregated information about iOS smartphone behavior (e.g., daily smartphone use or number of pick-ups for a particular app) but still lack the granularity of Android logging methods, which usually capture and timestamp every interaction with the device. Moreover, these procedures need intensive data processing protocols (i.e. for the processing of a video feed), and require effort, expertise and time from researchers to, for instance, tailor processing scripts to very specific use cases (Baumgartner et al., 2023). Given that platforms and their features and designs are constantly evolving (Bayer et al., 2020), these protocols also require frequent updating. Importantly, the process of providing daily screen captures is also rather cumbersome for participants, which may lead to attrition. In the study of Ohme et al. (2021), for instance, only 11.6% of the full sample ultimately donated their mobile log data. To avoid this, incentivization may need to increase, creating a higher financial burden on researchers.

In this paper, we present an alternative solution for leveraging data donations to capture screen data from iOS smartphones. The procedure that we present in this paper leverages high-quality, granular digital trace data of the participant's smartphone use, while requiring very little effort, budget, time and expertise from the researcher to implement, and demanding very little effort from the participant. To set everything up participants need to tick a box on each device they want to include (e.g., their Mac, iPhone, iPad, Apple Watch). Then, to donate the data they need to follow an easy step-by step process that takes less than 5 minutes. Because the procedure requires synchronization between an iOS, iPadOS or WatchOS device and a Mac computer, the use of it is conditional on the participant having a Mac computer and all other Apple devices should be linked to the same unique Apple ID. Nonetheless, we believe that our method offers additional ethical and legal benefits, and affords a novel way to meet the scientific goal of collecting the same granular use data from iPhones, iPads and Apple Watches that can be collected from Android phones.

As for the scientific goals, our procedure also opens new opportunities: Because it requires the synchronization with the user's Apple computer device, our procedure allows to capture data from *all* (Apple) devices that the user has embedded in their personal digital ecosystem under the same Apple ID. This has great value. After all, iPad data and laptop data are often neglected in current studies involving mobile sensing, while it is clear that cross-device usage is an important element of how people interact with the online world (Montanez et al., 2014; Tseng et al., 2023), and that people show different usage patterns on different digital devices (Bröhl et al., 2018). Since laptops, tablets, smartwatches, and smartphones can all be part of an individual's personal media ecology, excluding data from these devices when analyzing online behavior risks overlooking significant aspects of digital behavior, such as device switching.

In what follows, we first explain the protocol that we developed. Together with the manual in Section 8.3, researchers should be able to implement this protocol themselves. Next, we discuss the main benefits and drawbacks of our approach.

Method

Raw MacOS, iOS, iPadOS and WatchOS Data Donations as a Solution

Most current operating systems have a built-in app for usage logging. On iOS this application is called Apple Screen Time (on Android it is called Digital wellbeing). While the main goal of these built-in applications is to give smartphone users insight in their own usage (e.g., how often someone opens a certain app, how long they use their smartphone, how much notifications they receive), the applications can also be used by scholars to gain access to and thus measure participants' smartphone behavior without the need of installing a third party app through 'data donations' (Ohme et al., 2021). Data donations, such as those used by Baumgartner et al. (2023), shift the role of the user from someone whose data is collected to someone who collects their own data, and donates them to the researcher. In existent Screen Time data donation procedures (e.g. by using screenshots), the data that the user collects, is data that is already collected through the device or platform and has been pre-processed to populate the interface (e.g., Ohme et al., 2021). Our procedure accesses the 'raw data' as collected by the operating system itself and entails the parsing of that raw data, without any other aggregation and minimal processing.

A Step-by-Step Guide to Access MacOS, iOS, iPadOS and WatchOS Usage Data

To our knowledge, prior to 2019 there was no way to access the raw data of iOS Screen Time. From 2019 onward, however, Apple released Screen Time for the Mac and implemented a toggle to allow synchronization of Screen Time across Apple devices. With this toggle enabled, the Screen Time program on the Mac collects screentime and usage data of all Apple laptops, iPhones, iPads and Apple watches connected to the same Apple ID. In other words, local databases are produced that not only serve to inform the Screen Time visuals of the Mac, but that are also enriched with the usage data of all other devices from the same Apple ID for which this toggle is enabled. These databases are stored on the Mac of the (Apple ID) owner of the Apple devices. Our procedure uses these databases to get access to logged usage data of the Mac, iPhone, iPads and Apple Watches. The procedure is tested on a variety of devices running the latest software versions at the time of writing (iOS 26, MacOS Tahoe, iPadOS 26, WatchOS 26). Devices running older software versions are not supported because they store Screen Time data elsewhere.

In what follows, we provide a step-by-step overview of the procedure that participants can follow for donating their Screen Time data. An overview of each step can be found in Figure 6.1. The researcher is responsible for the data transfer and for parsing the data using the tool we have developed.

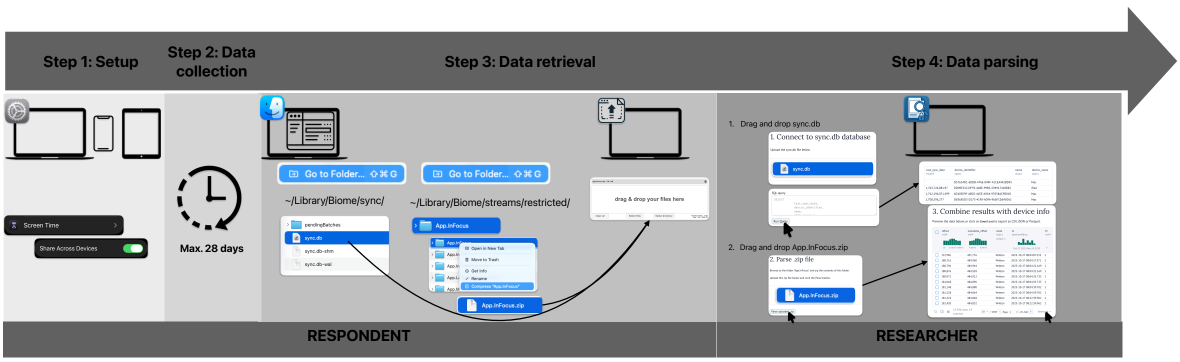


Figure 6.1: Overview of procedure.

6.3.b.a Step 1. Set-up: Toggle "Share Across Devices" on All Relevant Apple Devices

To collect Screen Time data from the participant's iPhone and/or other devices, the first step is to launch the Mac computer, and open the Screen Time tab in settings. In this tab, users can toggle the setting "share across devices". After this, they should toggle the same setting on their iPhone/iPad under the Screen Time settings tab. This setting needs to be enabled on each device they want to include in the dataset. The Macs should run at least MacOS 26, the iPads should run iPadOS 26, the iPhones should run iOS 26 and the Apple Watches should run WatchOS 26. The logdata for previous operating systems are located elsewhere. Hence, our current tool does not work on versions before these builds. This step should take place at the start of the data collection for the empirical study as any data from before the moment the toggle has been switched on will not be synchronized — notice that this is different from other programs (e.g., the app ActivityWatch can collect data from Mac computers (not iPhones/iPads) but immediately retrieves data from the past 10 days).

6.3.b.b Step 2. Data Collection

From the moment the switches are toggled, the usage data of the participant will be synchronized across all devices. The synchronization itself occurs through the personal iCloud of the respondent. Only the last 28 days of usage data is stored in the database, so for studies that run longer, participants need to donate their data on an ongoing, 4-week basis. Once the data collection phase has ended, the participants need to locate all necessary files. They will have to locate both (1) the identification file, to identify the different devices and (2) the Screen Time databases where all their Screen Time data is stored.

6.3.b.c Step 3. Data Retrieval: Compress and Donate the Folder Containing the Data

The identification file (sync.db) is an SQL database. In this SQL database, the devices can be identified using the *device_identifier* under the *DevicePeer* table. The column *platform* lists the type of device (1=iPadOS; 2=iOS; 3=MacOS; 4=MacOS; Apple watches are not included in this database), if the column *last_sync_date* is empty, the identifier is the one of the Mac on which the file is accessed.

The Screen Time data is stored in the App.InFocus folder. If different devices are synced (see step 1.), there will be two folders inside the App.InFocus folder, "local" and "remote". The "local" folder contains all the logdata of the machine on which you are accessing the folder (the Mac). The "remote" folder contains one (if only one other device is synced) or multiple subfolders named using different identifiers (e.g., DD09E32C-DF93-448E-99B3-594917418EB2). The databases inside these folders consist of SEGB encoded biome files which need to be parsed before they are usable. To do so, the parent folder App.InFocus should be compressed and donated to the researcher by the participant together with the identification file.

The identification file and the Screen Time databases are thus located on the filesystem of the Mac computer. At the time of writing this manuscript, this was under the pathnames:

- Identification file: *~/Library/Biome/sync/sync.db*
- Folder containing screentime data: *~/Library/Biome/streams/restricted/App.InFocus*

To collect the files from the respondents, the files will need to be transferred using a file transfer system that is provided by the researcher. For this step the researcher can use any service that is secure and falls under the legal and procedural regulations of their institution and country.

6.3.b.d Step 4. Data Parsing by Researcher: Process the Encoded Databases into Usable Format

To parse the data the researcher should run our developed tool. The tool does three important things: it parses the biome files into a usable format, it queries the sync.db file for device identifiers and it creates combined dataframes of the Screen Time data per device. To illustrate the procedure, we included the tool (which can be run locally through a docker image) and example files on the OSF page (https://osf.io/6puxm/?view_only=b44063a220204bdca33b964e2e8a4ebd).

The tool provides a Screen Time dataset for each linked device. The processed datasets are structured as follows: there are 7 different columns in the json/csv files necessary to interpret the usage data, see Table 6.1. The column ***timestamp*** refers to the timestamp of the log entry (e.g., 2025-10-30 09:03:48.338232). ***open_close*** logs whether an app is opened (1) or closed (0), each app usage will thus have two entries in the dataset (1 and 0). The usage time can be calculated by subtracting the timestamps of both entries. ***app_bundle*** shows the App Bundle ID (e.g. net.whatsapp.whatsapp) with which the respondent interacted. ***last_sync_date*** is the last time an entry was written to the database (UNIX timecode). This can help when different files should be merged. The first entry will never be more than 28 days before this last_sync_date. ***device_identifier*** is a unique identifier that indicates the device from which the data is logged (e.g. F5096A2A-ADAB-5981-B534-DE8DCC8EE313), if these fields are empty, the data is logged from the Mac from which the database is used. ***device_name*** (e.g. iPhone) is the linked name from sync.db. For each file there will be only one device name included as the datasets are divided per device.

Only apps in focus are logged. An app is in focus when the app is shown in the top left corner on the mac on an active screen (if multiple screens are used), for the iPhone and iPad this is the app that is visible or that is interacted with (if multiple apps are open), for the Apple Watch this is the app that is open and visible. Apps that are running in the background but are not visible or active are not logged.

Most of the time the App Bundle ID is easily linked to the specific app, however, in some cases the app name is not part of the App Bundle ID. The App Bundle ID for SnapChat, for example, is com.toyopagroup.picaboo. Prior to interpretation, the different application codes will need to be recoded into their respective apps. Overviews of often used App Bundle IDs can be discovered in online repositories[15]. However, not all App Bundle IDs will be included in these static repositories because of the dynamic nature of digital app usage.

Table 6.1: Example of the relevant columns included in the export of Screen Time Data.

timestamp	open_close	app_bundle	app_version	last_sync_date	device_identifier	device_name
2025-10-31 10:04:33	1	com.openai.chat	1.2025.294	1764327715	9CE1FB04	iPhone
2025-10-31 10:06:15	0	com.openai.chat	1.2025.294	1764327715	9CE1FB04	iPhone
2025-10-31 10:06:19	1	com.apple (...)	26.0.1	1764327715	9CE1FB04	iPhone
2025-10-31 10:06:36	0	com.apple (...)	26.0.1	1764327715	9CE1FB04	iPhone

[15] https://support.codeproof.com/apple-device-management/i-os-app-bundle-id

timestamp	open_ close	app_bundle	app_ver- sion	last_sync_ date	device_identifier	device_ name
2025-10-31 10:06:40	1	org.whisper (...)	7.82	1764327715	9CE1FB04	iPhone
2025-10-31 10:06:43	0	org.whisper (...)	7.82	1764327715	9CE1FB04	iPhone

6.3.b.e Step 5. Optional: End Data Collection

Unless explicitly specified otherwise by the participant, researchers can now inform the participant about how to end the data collection. As the procedure leverages built-in features of the iOS, MacOS and iPadOS ecosystems, the participant can stop the data collection by switching off the screen time feature. Nonetheless, as no data is being shared automatically, respondents can also keep on using the ScreenTime feature without sharing any data.

Discussion

Apple does not provide granular ScreenTime data through its ScreenTime APIs, nor is such data included when users request their personal information through the Data and Privacy portal. Consequently, obtaining usage data from Apple devices remains a significant challenge for both individuals and researchers.

In an earlier version, this study adopted a different strategy with the same idea: exploiting the comparatively less restrictive macOS environment as an access point. At that time, the relevant Screen Time information was stored in the KnowledgeC.db SQLite database, and extraction required only one straightforward SQL query. However, during the review phase, Apple released major operating system updates (iOS 26, iPadOS 26, macOS 26) that rendered this method ineffective. The data was migrated to less interpretable Biome SEGB files, which underpins the current methodology.

This development is a perfect example of the "moving target" problem, wherein Apple's gatekeeping role enables it to invalidate established data access techniques through routine system updates (Bayer et al., 2020). Although a workaround was identified, this experience underscores the fragility and transience of existing access mechanisms and the power imbalance between researchers and these companies. While current legislation efforts in Europe (e.g., GDPR, DSA) and in other regions in the world (e.g. state level legislation in USA, PIPL…) should enable researchers to access specific use data (through data donations or otherwise), current practices are still fragile and vulnerable to one-sided changes by the companies.

Benefits, Limitations and Risks

To our knowledge, this article is the first to describe the current procedure for collecting raw, behavioral digital use data of iOS (and iPadOS, WatchOS and MacOS) devices in a privacy sensitive and transparent way. By collecting only research-relevant data, we adhere to data minimization principles. A key benefit of our procedure is that participants only share safe encrypted data which are not easily decrypted without the right software.

A major caveat of our procedure, however, is that it can only be used when participants also have an Apple Mac device. This is especially relevant as people having access to a Mac and other Apple devices might meaningfully differ in sociodemographic profile and psychological traits. Moreover, it is only possible to collect 28 days of data through this method; for longer measurement periods, intermittent extractions must be organized, which can be cumbersome for both researchers and participants.

Finally, a possible risk as discussed above is that Apple changes the way they store, sync, code this dataset. This could possibly render the tools or step-by-step guide inoperative. Consequently, the moving target problem remains, and there are no guarantees that this procedure, which works today, will work in the future.

Nonetheless, despite these caveats, we believe the current procedure is less cumbersome and less effortful for both researchers and participants than alternative data donation procedures. For this reason, we consider our approach an important step forward towards integrating behavioral data from iOS and Mac users in research.

Open Practices Statement

Materials and tools are available at https://osf.io/6puxm/?view_only=b44063a220204bdca33b964e2e8a4ebd. None of the reported studies were preregistered.

Declarations

Funding: Funding support for this article was provided by the European Union's Horizon 2020 research and innovation programme under the European Research Council Starting Grant agreement 'DISCONNECT' No. 950635 and the Research Foundation Flanders (FWO-Vlaanderen) under Grant agreement 'Disconnect to Reconnect' No. S005923N and 'CostOf24/7' No. G075524N.

The funding body was not involved in the study design, the collection, analysis and interpretation of data, the writing of the report or the decision to submit the article for publication.

Conflicts of interest/Competing interest: The authors have no relevant financial or non-financial interests to disclose.

Ethics Approval: The study received ethics approval from the Institutional Review Board of Ghent University (ref. 2022-37).

Consent to participate: Not applicable.

Consent for publication: Not applicable.

Availability of data and materials: All materials are accessible through and downloadable through the OSF platform https://osf.io/6puxm/?view_only=b44063a220204bdca33b964e2e8a4ebd.

Code availability: All code is accessible through the OSF platform https://osf.io/6puxm/?view_only=b44063a220204bdca33b964e2e8a4ebd.

Discussion

This dissertation set out to examine the extent to which the theoretical interpretation, granularity, and scope of digital trace data shape what can be validly inferred about digital media use and well-being. The motivation for this endeavor was rooted in a broader paradigm shift within the field: as researchers increasingly move from retrospective self-reports of screen time toward passively collected behavioral data, the promise of greater objectivity and precision has been accompanied by a set of underexamined validity challenges: While log data circumvent recall bias and social desirability effects, they do not eliminate measurement problems. Moreover, the choices researchers make in turning raw event logs into analyzable constructs, from preprocessing and feature construction to device selection and interpretation of behavioral indicators, shape the conclusions that follow.

Across a methods chapter, three empirical chapters, and a tool chapter, this dissertation examined one or more sub-research questions, drawing from an intensive longitudinal dataset collected as part of project DISCONNECT, which combined experience sampling with passively logged smartphone and computer trace data.

First, the Methods chapter, **Chapter 2**, provided a detailed and explicit account of the trace-to-indicator pipeline underlying all empirical studies, documenting how decisions regarding data cleaning, event inference, time-window alignment, and feature construction are themselves sources of variation that shape the kind of theoretical claims that can be made. In doing so, the chapter aimed to contribute to ongoing discussions about validity, reproducibility, and transparency in digital trace data research.

Chapter 3 ("Connected Yet Cognitively Drained") addressed the first sub-research question, namely whether passively logged behavioral data can correspond with psychologically meaningful states. The study examined the state of online vigilance, and its momentary and lagged relation to cognitive fatigue. The online vigilance construct was assessed both via self-report and using digital trace data. Surprisingly, the findings revealed that the behavioral smartphone features designed to capture online vigilant behavior showed only weak associations with self-reported online vigilance. At the same time, this self-reported online vigilance was a consistent predictor of mental fatigue, whereas associations with the behavioral measure were absent. The chapter concluded that subjective experience appears to be more consequential than observed behavior in explaining the downstream well-being implications of digital connectivity, thus showing the interpretative limits of treating trace data as stand-ins for psychological states.

Chapter 4 ("Always On, Always Rushed for Time") shifted the focus to the second sub-research question, examining how granular and dynamic features of logged behavior relate to momentary well-being. By disaggregating smartphone use into four behavioral features (duration, frequency, fragmentation, and notifications) across four app categories (email, social media, chat, and work communication), the study demonstrated that category-specific and dynamic indicators revealed within-person associations with feeling rushed. Critically, these associations operated largely through a subjective mechanism, perceived juggling load, thereby also contributing to the first sub-research question by reinforcing that trace data require theoretical interpretation to become psychologically meaningful.

Chapter 5 ("Computers Matter") addressed the third sub-research question by including computer log data alongside smartphone logs for the first time. The chapter showed that adding computer data substantially altered screen time estimates (on average by +136%), and revealed frequent cross-device dynamics such as device switching (around 66 times per day), as well as overlapping use, and markedly changed the observed associations between digital media use and momentary and daily mood. Models using only smartphone data produced few and inconsistent associations, while computer and cross-device models showed more consistent patterns – particularly for work- and email-related activities. These findings demonstrate that measurement scope is not a minor methodological detail, but an important determinant of the inferences researchers can draw.

Chapter 6 ("Making the Impossible Possible") complemented the empirical work by presenting a novel procedure for collecting granular screen time data from iOS devices – a platform that remained largely inaccessible to researchers due to Apple's restrictive policies. This procedure also includes a tool we developed which does some of the "heavy lifting": it parses the raw biome files into a usable format, queries the sync database for device identifiers, and creates combined data frames of Screen Time data per device which can be used for subsequent feature creation. The tool is offered in two deployment modes: researchers can either access a publicly hosted Marimo Notebook online, or run a containerized local instance through Docker. By providing both options, the tool accommodates varying institutional constraints: from researchers who need data to remain on local infrastructure for privacy or ethical reasons, to those who prefer the convenience of a browser-based workflow. Moreover, both the tool and all accompanying materials are shared openly via the Open Science Framework, so that the data processing pipeline becomes open and reusable by the broader research community.

Together, these chapters arrive at a central insight: the shift from self-reported to logged measures of digital behavior does not in itself resolve the measurement challenges facing digital well-being research. The validity of what we can conclude about digital media use and well-being depends not only on *whether* we use trace data, but on *how* we select and interpret behavioral indicators, at what level of granularity we operationalize them, and from which devices we collect them. The sections that follow discuss the implications of these findings, first for the use of log data in digital well-being research specifically, then for broader debates within the field, before turning to limitations and societal implications.

Contributions of the Use of Log Data in Digital Well-Being Research

Interpretative Gap

An overarching finding is that logged behavioral indicators, on their own, provide an incomplete account of the relevant cognitive-behavioral states they supposedly represent. Behavioral indicators capture *when*, *how long*, and *how often* someone interacted with a device, but not *how* these interactions are truly experienced by the user. More specifically:

In Chapter 3, we attempted to construct theory-driven behavioral features of online vigilance, i.e., by looking at monitoring sessions, notification reaction speed and responsiveness, which mapped to two specific dimensions of the construct. Yet, the measures showed only weak associations with self-reported online vigilance, and also performed poorly in predicting mental fatigue, while self-reported online vigilance was a robust predictor.

Chapter 4 did present meaningful associations between smartphone use and feeling rushed, but mostly operating through a subjective mechanism. Specifically, the link between granular smartphone usage features and feeling rushed was mediated by perceived juggling load. The subjective appraisal of having too many tasks to manage across roles served as a bridge between behavioral data and well-being outcomes.

Combined, these two chapters illustrate an interpretative gap, whereby the subjective experience appeared to trump behavior in explaining the implications digital media use. This brings the question: were the specific features chosen as indicators of the underlying state poor behavioral representations, or should we accept that self-reported subjective experiences simply trump behavioral measures in explaining the implications digital media use? A growing body of work indeed points towards subjective beliefs, perceptions and evaluations of digital media use as playing a central role in shaping well-being outcomes. For instance, Ernala et al. (2022) demonstrated subjective *moderation*, i.e., the same behavior related differently to well-being based on the users' beliefs; Lee & Hancock (2024) showed that *mindsets* explained more variance in well-being than behavioral use measures; a finding that Parry & Coetzee (2025) replicated (albeit finding some differences to the original work of Lee and Hancock).

Alternatively, this gap is not an artefact of one particular operationalization or construct, but perhaps a more general feature of the relationship between logged behavior and subjective well-being. Therefore, the question for the broader field of digital wellbeing research is not *whether* to use either self-reported or behavioral data, but *how* do we create research designs that can integrate both in theoretically productive ways? Such a path forward is presented in the recent social constructivist framework of Wolfers (2024). This framework considers perceptions of media use not as biased proxies of objective behavior, but conceptually distinct constructs that are themselves socially constructed and independently consequential for mental health. This dissertation's empirical findings are broadly consistent with this emerging theoretical direction: well-being implications of digital behavior appear to be filtered through different subjective processes that are not reducible to the behavior itself.

The interpretative gap, however, operates not just at the level of the user; it also runs through the research process itself, as we have shown in our Methods chapter (Chapter 2). Each of the steps taken in this process involves the researcher's judgement about what constitutes relevant behavior. This means that trace data occupy a position between objective and subjective. This dual subjectivity, in how users experience their behavior and how researchers construct behavioral indicators, makes it difficult to label trace data as (completely) objective. Rather than seeing these traces as neutral records, and self-reports as biased, the field should acknowledge that both data sources require interpretation, albeit of different kinds.

Apart from the necessity to document decisions fully (see below for a discussion of the implications in relation to Open Science), this interpretative gap leads to two broader implications. First, for research design: if subjective mechanisms are essential for understanding the well-being implications of digital behavior, then linkage designs combining trace data with ESM are not merely a *nice-to-have*, but more of a necessity. Studies relying on trace data alone risk producing findings that are perhaps precise in their behavioral description, but lacking in psychological explanation.

Second, for theory: the findings raise the question of what trace data then *are* indicators of. If they are not reliably capturing psychological states such as online vigilance, and if their associations with well-being depend more on subjective mediators, then how do they contribute independently? Trace data are perhaps best understood as indicators of behavioral *structure* that become meaningful in research when interpreted through theory and triangulated with subjective data.

In sum, the dissertation's findings suggest that trace data are best understood as capturing a different level of description – the behavioral-structural level – that requires integration with subjective data to become psychologically meaningful, not as a more accurate version of what self-reports measure or as a fully independent source of information. This has consequences for the methodological debate, where the central question should be: at which level of description does each data source operate, and how should they be combined?

A more practical implication of working with these data is that investing in ever more fine-grained logging could yield diminishing returns unless it can be matched with equally fine-grained subjective data. Increasing behavioral precision without attempting the same for psychological precision would, in other words, not help close the interpretative gap.

Granularity

A recurrent critique of existing research on digital media use and well-being is its reliance on undifferentiated measures of screen time, typically total duration or total frequency of device use aggregated at a daily level (Kaye et al., 2020; Meier & Reinecke, 2021). As discussed in the introduction, treating screen time as a monolithic construct clashes with theoretical models that conceptualize digital well-being as the product of dynamic interactions between person-, context-, and device-specific factors (M. M. P. Vanden Abeele, 2021). If what matters is not just how much people use their device, but what they do, in which temporal pattern, and in which context, then aggregate indicators may flatten the variation the theory predicts to be relevant. This dissertation treated the explanatory value of increased granularity as an empirical question. Across the empirical chapters, the findings suggest that granularity does matter, but not unconditionally, and not in a straightforward "more is better" manner.

Chapter 4 provides the clearest test of this proposition. By disaggregating smartphone use in four behavioral features across four app categories, we were able to examine a matrix of category × feature combinations. This disaggregation revealed associations that aggregate measures would have obscured. For instance, email and work communication features showed the strongest association with feeling rushed, while social media features showed no significant direct associations. Had we relied on total smartphone duration alone, these differential patterns would have collapsed into a single, likely weak and uninformative, coefficient. In this sense, granularity proved analytically informative: it allowed us to identify which kinds of digital behavior related to a specific well-being outcome, and enabled to begin distinguishing functional contexts of use.

This categorical decomposition represents a shift from studying screen time to studying screen use, where the functional character of digital behavior differs across categories which map onto distinct theoretical expectations (Meier & Reinecke, 2021). Email and work communication apps, for instance, carry connotations of professional obligations and task demands, whereas social media use may serve as leisure. Treating both as interchangeable units of screen time obscures these functional differences. Notably, Chapter 5 extended this observation by showing that even within the same category, the device on which an activity occurred shaped its association with mood: email use on the computer was consistently linked with lower happiness and higher feelings of being down, while email use on the smartphone showed weaker or non-significant associations. This implies that categorical granularity alone is not sufficient; the device context further explains what a given category of use means in practice.

Beyond the question of what people do, this dissertation also explores how behavior is distributed in time. The fragmentation feature, capturing whether screen time accumulated through sustained sessions or frequent short bursts, proved to be a relevant predictor in both Chapters 4 and 5. In Chapter 4, fragmented smartphone use was positively associated with feeling rushed, with this association operating mostly through perceived juggling load. In Chapter 5, multiple smartphone, computer, and combined-device fragmentation patterns showed positive associations with feeling down, and negative associations with feeling happy. These findings suggest that the temporal structure of digital behavior carries information that volume measures alone do not capture.

Yet, it would be misleading to conclude that finer granularity automatically produces stronger or more coherent findings. In Chapter 4, effect sizes for granular within-person associations remained small, even when category-specific and temporal features were employed. Chapter 5 showed that not all granular features outperformed their aggregate counterparts: some category-specific models did not yield significant associations while more aggregated ones did. Granularity adds analytical resolution, but more resolution does not guarantee more signal. It can also mean more noise and a greater risk of theoretically incoherent patterns. Consequently, deciding which kind of granularity is most theoretically relevant given the research question is key.

Finally, granularity raises a practical tension that should be acknowledged. Finer-grained operationalization demands more behavioral data, preprocessing decisions, and more analytical complexity. Each of these introduces additional researcher degrees of freedom, but also new threats to reproducibility. As we have shown in the Methods chapter, the computation of these features requires multiple decisions. An increase in more granular features should therefore also be accompanied by fuller documentation practices, allowing for evaluation and reproduction of these operationalizations. Else, this risks producing findings which are more complex, but not more trustworthy.

Scope

A third contribution of this dissertation concerns the scope of behavioral measurement. While digital well-being research has increasingly embraced passive smartphone logging, it has largely treated the smartphone as a sufficient proxy for individuals' broader digital engagement. The findings of Chapter 5 demonstrate that this assumption is flawed.

First, restricting measurement to smartphones results in substantial construct underrepresentation. In our sample, including computer log data increased average daily screen time estimates by roughly 136%. This is not a small correction, but could signal a larger shift in how digital engagement is quantified. Participants with near-identical smartphone use displayed very different overall screen time patterns once computer use was included. Thus, reliance on a single device both underestimates usage and misrepresents individuals' behavioral exposure.

Second, limiting measurement scope obscures qualitatively distinct behavioral dynamics. Multi-device logging revealed frequent device switching, overlapping use, and context transitions. These patterns, which smartphone data alone would not allow us to infer, were both common and associated with lower well-being. In this sense, scope affects which theoretical mechanisms can be tested, not only how much behavior is measured. Processes such as attentional fragmentation or media multitasking would not be visible in single-device designs.

Third, device scope shapes inferential conclusions. Smartphone-only models yielded few and inconsistent associations with mood, whereas computer and cross-device models produced clearer patterns, particularly for work- and email-related activities. This indicates that, especially among adult populations, conclusions about digital well-being depend on what parts of their digital behavior are captured.

These findings relate to broader concerns about tracking undercoverage in digital trace research, where incomplete device capture leads to systematic bias (Bosch et al., 2025). In the context of digital well-being research, undercoverage risks misestimating both exposure and its psychological consequences. Device selection therefore constitutes a theoretical boundary decision: it determines what counts as digital media use in the first place.

Chapter 5 of this dissertation demonstrates that multi-device logging is a feasible methodological direction, and by introducing a procedure for granular iOS data extraction along with open-source computer logging tools, Chapter 6 provides a practical pathway towards expanding measurement scope across devices.

Even the multi-device designs adopted in Chapter 5 capture only a partial view of participants' full device ecologies. Our data encompasses smartphones and computers, but leaves out, for instance tablets, TVs, gaming consoles, and wearables – devices that, for some, may constitute a non-trivial amount of daily digital engagement. Furthermore, the measurement foundation across the two devices was not symmetric. Certain behavioral indicators that proved relevant in Chapters 3 and 4, most notably notifications, were available exclusively from the smartphone, as the computer-based logging tool did not capture equivalent events.

The scope argument advanced here is not a claim that our two-device configuration is sufficient to capture a person's full digital ecology. Smartphones and computers represent only part of a broader set of digital devices that may include many others. Each additional device introduces its own data structures, logging conventions, technical limitations and constraints that complicate integration into a coherent measurement framework. Fully capturing an entire device ecology thus remains an aspiration that current tools cannot achieve. What Chapter 5 demonstrates is that expanding beyond a single device already meaningfully changes what we observe. There is little reason to assume that further expansions to other, additional, devices would not produce similar shifts. However, as we discussed before, expanding studies to include more devices and even richer data will not automatically close the distance between observed behavior and psychological experience. Investing in broader device coverage therefore risks increasing measurement complexity without adding explanatory value, unless it is accompanied by increased efforts to interpret what these additional behavioral records mean for the individuals generating them. In other words, expanding scope is most productive when contextual and subjective data can provide meaning to the behavioral patterns and the psychological processes they are meant to explain.

Transparency and Reproducibility

The methodological shift towards trace data was motivated, in part, by the inherent validity problems that come with self-report data. Logged behavioral data offer the promise of greater objectivity. Across this dissertation, and specifically in the Methods chapter, however, we have come to see this promise is incomplete.

While trace data do eliminate recall bias and social desirability effects, the trace-to-indicator workflow documented in Chapter 2 reveals that the transformation of raw device logs into analytical variables involves numerous consequential decisions – about what to retain, how to parse events, where to set thresholds, and how to operationalize features – each of which introduces researcher judgement into the data. Because these decisions cascade through the pipeline, with early-stage choices impacting subsequent indicators, they are not just isolated technical operations, but sources of variation that shape what the data can and cannot show. Providing transparency about these decisions is therefore key to increasing reproducibility of log-based research.

The pipeline document in Chapter 2 makes visible a set of decisions that are rarely reported in sufficient detail for others to evaluate or reproduce. This raises a broader question about where such pre-analysis processing sits within current thinking on computational reproducibility. Existing frameworks have articulated the importance of code sharing, environment encapsulation, and tiered verification (Bleier, 2025; Schoch et al., 2023), but have, understandably, focused more on the analysis stage of the research pipeline, where statistical code operates on data that is already in its final, analyzable form. The pre-analysis processing decisions that are particularly consequential for trace data have received comparatively less attention. The detailed account offered in Chapter 2 is a response to calls for such documentation (Baumgartner et al., 2023; Parry & Klingelhoefer, 2025; Parry & Toth, 2025), but transparency alone does not fully resolve the problem. Even a thoroughly documented pipeline represents a single path chosen from a range of alternatives. Complementary approaches, such as multiverse-style sensitivity analyses which could be used to assess how different processing decisions affect downstream conclusions (e.g., Murphy et al., 2025), or interactive computational documents (e.g., R Markdown, Quarto) that allow readers to adjust pipeline parameters and observe resulting changes, could further strengthen the reproducibility and evaluability of trace-based research.

Contributions to Broader Discussions in the Field

Above, I have provided a general synthesis of the findings of my empirical work, and have formulated an overarching response to the three core research questions guiding it. In the current section, I explore some broader implications of the study findings for the field of research on digital well-being, and digital media effects more generally.

Small Effects

A recurring finding across the empirical chapters is that effect sizes of behavioral indicators are small. This finding is consistent with the broader literature where small effect sizes have been reported across studies employing logged behavioral data and ESM designs (e.g., de Segovia Vicente et al., 2024; Elmer et al., 2025; Siebers et al., 2024). Several considerations could help interpret these small effect sizes. I want to discuss three of these considerations, which are also limitations of the current dissertation, in greater depth: (1) the potential for momentary effects to accumulate over time, (2) the possibility that effects have already stabilized, and (3) the sensitivity of effect estimates to the inclusion of control variables.

First, as we investigated momentary associations, an important question is whether these small effects might accumulate into a larger, more consequential impact. As we will also discuss in the limitations (see below), we put forward this accumulation-hypothesis when investigating online vigilance (Chapter 3), as we found both contemporaneous and lagged effects of online vigilance on mental fatigue. If these small effects could linger throughout the day and are not fully regulated, they could potentially accumulate over the course of the day. Thomas (2022) offers a systematic methodological framework for analyzing such dynamics, distinguishing between different patterns in the appearance and duration of media effects. Their typology includes, among others, the “drip” scenario, a pattern in which a single media exposure produces weak or absent effects, but the influence across repeated exposures can accumulate, leading to a meaningful overall effect. Modelling the effects observed in our study using such a drip framework could have provided more insight into whether momentary associations with fatigue build up during the day. At the time of performing the study, I was not familiar enough with these models to confidently employ them, but believe the observation of a t-2 lagged effect of online vigilance warrants such a future exploration.

Second, a complementary perspective, advanced by Baumgartner (2025), proposes that media effects may stabilize after repeated exposure. Drawing on theories of habituation, adaptation, and learning, Baumgartner argues that media effects are likely to diminish and eventually plateau as users grow accustomed to their media environments. If this is the case, effects are only empirically observable during an “effect-sensitive period”. For more established media users, such as the adult participants in our study, this would mean that the period during which digital behavior most strongly affected well-being outcomes may have already passed before data collection began. This could offer an explanation as to why our momentary associations remained small, although statistically significant. I think future research should consider how research designs could be developed to capture media habits when they start to form, or when changes in digital media use may be observed, for instance as a result of major life changes.

Finally, the inclusion of control variables when modelling media effects can make a noticeable difference to the effect sizes observed. Ferguson (2026), for instance, demonstrated that bivariate correlations between self-reported social media screen time and various well-being indicators largely disappeared once trait-level controls were included. This suggests that small bivariate associations might be artefacts, reflecting shared underlying factors, instead of direct media effects. A similar dynamic was observed in our own work: in Chapter 4, we initially modeled the association between smartphone use and time pressure controlling only for whether or not a participant had worked. Reviewers argued this did not sufficiently account for potential role blurring between work and personal life, hence an additional control for the time of day was added, which altered several of the observed effect sizes.

Relatedly, across the empirical chapters, behavioral features – such as duration, frequency, and fragmentation – were typically entered in separate models instead of being included together. Whether the small individual effects reported reflect overlapping or distinct variance in digital well-being outcomes remains an open question, and future research could benefit from modelling multiple behavioral features concurrently to clarify their relative and combined contributions.

Heterogeneity, Social Structure, Person-Specificity

An underexplored contribution of digital trace data is their capacity to reveal usage patterns structured by social roles and broader life circumstances. In Chapter 4, we offer some illustration of this: several of the associations between mobile communication features and feeling rushed or perceived juggling load were moderated by age, gender, parenthood status, and segmentation preference. For instance, participants with children showed stronger associations between email notifications and feeling rushed, while older participants showed weaker associations between chat notifications and this indicator of experienced time pressure. These moderations, while modest, hint at the fact that digital behavior is embedded in the demands and rhythms of people's broader daily lives: their work obligations, caregiving responsibilities, and role boundaries.

That said, there are several theoretical frameworks that articulate this argument more explicitly. M. M. P. Vanden Abeele (2021) situates digital well-being within a dynamic system where context-specific factors, including social roles and availability norms, shape how device use relates to well-being. Their model also points to processes of acceleration (Rosa, 2013; Wajcman, 2014) and the commodification of attention (J. Williams, 2018) as broader socio-cultural forces structuring this system. In Wolfers' (2024) social constructivist perspective on media effects, individual and situational factors are also addressed, including role-based contexts such as parenting or work, which according to Wolfers' model affect both objective media use and the perceptions of that use, while these perceptions themselves have consequences for well-being. Our Chapter 4 finding on the mediation role of juggling load provides some empirical grounding for these theoretical calls: mobile communication appears to intensify the collision of social roles, a point that has also been amply illustrated by my colleague Sara Van Bruyssel in her ethnographic dissertation work (Van Bruyssel et al., 2026).

Chapter 5 extends this point to the level of device ecosystems. The device types participants chose to include in the study ranged from gaming desktops and private laptops to dedicated work computers, meaning that the scope of what is captured by log data varies across individuals and is likely to co-vary with employment status or profession. This heterogeneity in device ecologies suggests that what constitutes someone's total digital engagement is itself socially structured. Chapter 5 argues that addressing this complexity may necessitate a shift towards n = 1 modelling perspectives, where each participant's unique device ecology could be modelled in relation to their own affective dynamics. From the measurement side, this point is reinforced by Keusch et al. (2022), who demonstrated that smartphone handling behaviors that act as barriers to sensor-based measurement are themselves sociodemographically patterned. Women, older adults, and less frequent smartphone users were more likely to exhibit these behaviors, meaning that trace data can be incomplete in ways that co-vary with the social positions that may also shape how digital behavior relates to well-being.

This dissertation does not make a major sociological contribution in its attempt to demonstrate how sociodemographic and contextual factors impact heterogeneity in associations between digital behavior and well-being. The sociodemographic moderators were exploratory, not theory-driven, and the self-selected sample, consisting of predominantly highly motivated adult knowledge workers, limited the extent to which variation could be examined. The question of how broader structures of inequality, gendered divisions of labor, or institutional norms around connectivity shape daily digital practices captured by trace data remains largely open, and points in a clear direction for future work.

Limitations

Partial Utilization of the Dataset

A major limitation of this dissertation concerns its single-wave scope. The broader ERC DISCONNECT project employed a multi-wave burst design with three separate data collection waves, yet this dissertation draws solely from Wave 1 data. This means that longer-term longitudinal dynamics across waves, including the potential to examine whether associations between digital behavior and well-being change, stabilize or build up over time – dynamics which recent theoretical work has argued are central to understanding media effects (Baumgartner, 2025; Thomas, 2022) – remain unexplored. This means that a core strength of the project's design has not yet been fully realized within this dissertation.

That said, the choice to focus on Wave 1 data was a deliberate and pragmatic one. As documented in the Methods chapter, the processing pipeline for combining Android trace data, computer trace data and mobile experience sampling, even within a single wave, was substantial. The sheer volume of data already required considerable investment before any meaningful analysis could take place. This type of work – cleaning, parsing, debugging – constitutes a form of hidden labor that is essential to the success of trace-based research, yet remains largely invisible in the final publications it enables, rarely recognized by current academic incentive structures (Parry & Klingelhoefer, 2025).

In addition, between waves, changes were introduced in the tools used for data collection, including the app used for experience sampling. By Wave 2, a new version of the passive logging app for Android users was released, which included an ESM module. This had consequences for both the ESM protocol and, to an extent, for the structure of the raw trace data. These changes, which were beneficial from the perspective of (Android) participants, meant that cross-wave alignment of trace data and within-wave alignment of ESM data were not reducible to simply appending datasets. This required additional harmonization steps, once again introducing further decision points and potential sources of error.

Data harmonization and alignment steps were carried out, albeit after Wave 3 data collection was completed. As such, these steps fell outside of the scope and timing of the current dissertation. My hope is that we can share these data with the broader scientific community, allowing other researchers to benefit from our data collection efforts and this unique dataset. This ambition responds to growing calls in computational social science to move beyond data-available-upon-request claims, which often go unfulfilled over time (Schoch et al., 2023), and toward more structured archiving practices that support third-party verification and cumulative science progress (Bleier, 2025).

Finally, even had cross-wave analysis been conducted, doing so would have introduced a new set of methodological challenges related to sample fragmentation. Across three waves, not all participants contributed to every wave: some participants from Wave 1 did not return for subsequent waves, while new participants joined in later waves. The combination of participant attrition and new recruitment means that any longitudinal analysis, spanning multiple waves, would necessarily operate on a fragmented sample, one in which the subset shrinks as more waves are required to overlap. This fragmentation is not unique to the cross-wave dimension, however; as the next section discusses, it also occurs in other forms within the empirical chapters of this dissertation.

Fragmented Samples

All three empirical chapters draw from the same Wave 1 data collection, but each chapter effectively operates on a different analytic sample. Our full sample consists of 1,315 participants, but only 774 provided usable Android trace data, and of those, only 106 also provided computer trace data. Chapters 3 and 4, which rely on smartphone traces alone, draw on a much larger subsample than the subsample used in Chapter 5, where both smartphone and computer trace data were required. This means that findings across chapters are not strictly comparable, as each is based on a different subset of the participant pool, and as data requirements become more stringent, selectivity likely played a larger role. Further, this selectivity might not be random: participants who provided both smartphone and computer data likely differ from those who only provided smartphone data (e.g., in terms of motivation, digital literacy, device ecology), echoing Bosch et al.'s (2025) point that undercoverage is shaped by participant characteristics.

A different, more granular, form of fragmentation arose within each chapter, particularly in Chapters 4 and 5, where category-specific features (email, social media, chat, and work [communication]) were modelled. The issue here concerns the interpretation and handling of data absence: when a particular participant showed no category-specific use during a given ESM time window, it is ambiguous whether this reflects genuine non-use (a meaningful zero) or data absence (a missing that should be excluded). If a participant *never* uses email apps on their smartphone across the entire study period, should their email duration be coded as a zero – implying that email non-use is a meaningful behavioral state to include in models – or should they be excluded from email-specific analyses entirely? This decision determines whether the analytic sample for a given app category includes only users of that category or the full sample. To our knowledge, no established convention exists in the field for resolving this ambiguity, and different choices would produce different samples and different results, connecting back to our discussion on researchers' degrees of freedom (Methods chapter). This ambiguity is impacted by the inherent ambiguity in app categorization itself, where decisions about how to label applications (using existing or custom labeling schemes) can alter which participants register any use in a given category (see also Elmer et al., 2025; Schoedel et al., 2022).

Another point related to missing data is the notion that, in our study on digital well-being, the behavior of interest – disconnecting from digital devices – may produce theoretically relevant missing data in both trace and ESM data streams. If a participant deliberately goes offline or shuts down their device, they will not receive or respond to their ESM prompts and will not generate log data via their smartphone. In other words, well-being during disconnection periods goes unrecorded. Importantly, the behaviors that produce this data loss are not marginal: Keusch et al. (2022) found that substantial proportions of smartphone users reported regularly leaving their device at home or not carrying it on their person, behaviors which would suppress both trace data and ESM data in a study like ours. This creates a non-random form of missingness: the moments where participants are most disconnected are also the moments least likely to be captured by the study design. In Chapter 3, we explicitly discuss this limitation, as several participants reported in the final questionnaire that their missing ESM values were often due to deliberate disconnection decisions. This observation, also noted by Klingelhoefer et al. (2024), reveals an intrinsic tension in ESM research on digital disconnection: the method intervenes on the reality it aims to capture, as the smartphone notifications required to collect ESM data are also a form of digital demand on participants.

Reactivity

The interdependence between study design and the behavior under investigation extends beyond the problem of missing data. A related concern is reactivity: the possibility that participants alter their digital behavior precisely because they know it is being monitored. In studies that combine passive logging with experience sampling, as ours does, reactivity can be present in both data streams and on different timescales. Toth et al. (2025) disentangled these two sources in their study. They found that smartphone logging induced small reductions in use frequency and duration that were most pronounced during the first five days before fading. This is consistent with earlier, smaller-scale evidence of habituation to monitoring (Toth & Trifonova, 2021). ESM prompts, by contrast, produced short-term increases in smartphone use lasting approximately two to three hours after each survey, likely because the act of picking up the phone to complete a questionnaire served as a connection cue that triggered further device use. This latter finding is particularly relevant to our design, where participants received six ESM prompts per day via their smartphone. On the ESM side, there is an additional, more conceptual, form of reactivity: our study of online vigilance (Chapter 3) required participants to attend to incoming notifications from the ESM app, effectively asking them to be vigilant toward their smartphone in order to report on their vigilance toward their smartphone. As we noted in that chapter, some participants explicitly remarked on this irony, indicating that the ESM notifications themselves became part of the attentional demands the study aimed to measure. Together, these observations suggest that our study likely captured behavior that was shaped, in part, by the measurement process itself. The existing evidence offers some reassurance that logging-induced reactivity tends to diminish within the first days of a study period, and that ESM-induced reactivity, while recurrent, is not long-lasting (Toth et al., 2025; Toth & Trifonova, 2021). Nevertheless, the possibility that both forms of reactivity shaped the behavioral data in our empirical chapters cannot be fully ruled out, and remains a source of uncertainty common to all intensive longitudinal designs that combine experience sampling with passive sensing.

Societal Implications

Screen Time Discourse and Public Policy

While this dissertation has primarily adopted a methodological lens, its findings are not without broader societal relevance. The methodological choices examined throughout these chapters directly have consequences for how digital well-being is understood and acted upon in public discourse and policy.

Public debate on screen time and well-being overwhelmingly treats screen time as a single number: hours per day. As we have discussed, this reflects some of the "conceptual and methodological mayhem" related to screen time (Kaye et al., 2020); the concept is underspecified, but central in popular concern and legislative action. Although research on the topic has increasingly moved away from self-reported and aggregated duration measures, this shift has not yet reached the evidence base that informs policy.

As an example, take the topic of smartphone bans in schools, which has received considerable policy attention in recent years. The evidence base informing these policies, however, remains limited, and critically, continues to rely on the very measurement approaches this dissertation has called into question. Recent work on the consequences of these bans in the UK (Goodyear et al., 2025) and the Netherlands (Vanluydt et al., 2026) has concluded that there is little to no impact on youth's screen time or well-being, leading others to state that these bans are unproven or have failed (Ferguson, 2025, 2026). In each of these cases, the measurement of smartphone use is limited to what students self-report. In some cases, this involved prompted self-reports, where students were asked to consult iOS Screen Time or Android Digital Wellbeing dashboards; in others, screen time was captured using a single survey item asking students to estimate daily hours across all devices. As this dissertation has demonstrated, self-reported screen time differs systematically from logged behavior (Parry et al., 2021), but also obscures granularity and usage patterns that may be more relevant than just aggregated duration alone. For instance, the question of whether these bans might only shift or redistribute screen time towards other periods outside school hours, or towards other devices such as laptops present in classrooms, has already been raised (Vanluydt et al., 2026; Weiss & Bonell, 2025). Until log-data approaches find their way into policy-relevant research, the evidence base for interventions such as smartphone bans will remain constrained by the measurement limitations this dissertation has sought to address.

Digital Divide(s) in Research Participation

The methodological advances in digital behavior tracking discussed in the previous sections have consequences that are less frequently discussed: they can produce and reproduce digital divides that shape *who* is studied and *who* can conduct these studies. boyd & Crawford (2012) warned that unequal access to Big Data creates new stratifications in research, between the "Big Data rich" and "Big Data poor", and that these inequalities bias both the data and the kinds of research that follow. In trace-based digital well-being research, this claim holds.

A first source of bias concerns operating system constraints. The majority of smartphone logging research relies on Android devices, as Android has historically afforded researchers access to detailed, event-level usage logs (Parry & Klingelhoefer, 2025), while iOS imposes far more restrictive data access policies. Existing workarounds, such as screen capture donations of Apple Screen Time (Baumgartner et al., 2023; Ohme et al., 2021), yield only aggregated data and can impose a higher burden on participants, possibly leading to attrition. The procedure presented in Chapter 6, which extracts raw iOS Screen Time data, is a promising step but requires participants to own both an iPhone and a Mac, which restricts eligibility to users embedded in the Apple ecosystem.

Beyond platform constraints, participant self-selection further narrows who is represented. Trace data collection typically requires installing software, granting extensive permissions, and sustaining that participation over time – conditions that favor individuals with higher digital literacy and lower privacy concerns, and higher trust in researchers (Keusch et al., 2024; Pankowska et al., 2025). The DISCONNECT project illustrates this concretely: our recruitment in collaboration with De Standaard yielded a sample where 93% held at least a bachelor's degree; participation was incentivized through personalized well-being feedback without offering financial compensation, likely attracting people already motivated to reflect on their digital habits; and of 1,315 ESM participants, around 70% were Android users. Each selection stage progressively narrows the sample towards a specific profile: educated, digitally reflective, Android-using adults. The risk, then, is that our understanding of digital well-being is built on the experiences of a particular subset of the population. Whether these findings generalize to those who are less digitally literate, who use different devices, or who lack the motivation or time for intensive longitudinal studies remains an open question.

The divide extends equally to the researcher side. Working with behavioral trace data requires expertise in data collection design, processing of complex log files, and computational analysis. These skills are not uniformly distributed across the research community. As boyd & Crawford (2012) observed, computational skills are generally restricted to those with a technical background, raising the question of who is advantaged in research contexts that increasingly require these skills.

In practice, even the data access methods themselves are fragile. As we outlined in Chapter 6, during the development of our original iOS logging tool, an Apple operating system update rendered our data extraction procedure inoperable mid-study, illustrating the power imbalance between researchers and platform companies. Parry & Klingelhoefer (2025) note more broadly that commercial organizations frequently act as gatekeepers, mediating and restricting access to trace data, and that the specialized infrastructure required can present a significant barrier for researchers without institutional support.

Reproducibility can also create knowledge gaps in, for instance, programming practices among computational social scientists, many of whom are self-taught and lack formal training in software engineering (Schoch et al., 2023). Proprietary tools and APIs impose additional economic barriers, and heavy computational resource requirements can render research effectively irreproducible for those without access (Bleier, 2025; Schoch et al., 2023).

These dynamics risk reinforcing the hierarchies boyd & Crawford (2012) cautioned against: well-resourced research groups with computational infrastructure and skills are best positioned to conduct trace-data research, while those at small institutions or in less computationally oriented disciplines are excluded from these types of research designs. If we want to understand digital well-being across diverse populations, then we must remain critical of the exclusionary nature of these dynamics.

Concluding Remarks

This dissertation began from a simple premise: that the move from self-reported screen time to passively logged behavior promised more objective and precise measurement of digital media use, and thereby clearer answers about its relation to well-being. The work presented here suggests that this promise is real but conditional. Trace data do circumvent the recall and desirability problems of self-report, and they open analytical possibilities (e.g., category- and feature-level granularity, cross-device scope, temporal dynamics) that aggregated self-reports cannot. Yet across the empirical chapters, the same lesson recurred in different forms: the validity of what we can conclude does not follow automatically from the act of logging behavior. It depends on how indicators are constructed, at what granularity they are operationalized, from which devices they are drawn, and on whether they are interpreted alongside the subjective experiences that give behavior its meaning. The take-home message, then, is not that trace data are unreliable, nor that self-reports should be reinstated, but that neither data source is self-sufficient. I hope this dissertation can guide others in designing studies that combine them deliberately, document the choices they make in connecting raw trace data to analytical claims, and remain honest about the populations and devices their designs can and cannot reach. If this dissertation has a forward-looking contribution, it is an invitation for others to treat the integration of behavioral and subjective data—and the transparency of the pipeline that links them—not as a methodological afterthought, but as a central design problem of the field.

Appendices

Appendix A — Supplementary Material for Chapter 3

Participants

Of the 3,065 participants that originally signed up for participation, 1,449 (47.3%) responded to the onboarding invitation, provided informed consent and began receiving questionnaires. After removing participants that did not meet our eligibility criteria or who completed too few ESM questionnaires (pre-registered minimum of 8), the data of 1,315 participants were retained for formal analyses.

Persons willing to sign up for our study were first directed to a webpage that provided them with a study overview, and then were directed to complete an online form that verified they met our eligibility criteria: that they were aged 18 years or over, owned a smartphone, and were willing to complete the two-week ESM phase. After signing up, participants immediately received an automated email that provided full study information. Up to one week afterwards, participants received an invitation email which contained a link that directed them to our onboarding website. Here, they once more found all study information (presented both in print and in a video format) as well as detailed instructions to download and install one or two smartphone apps. Participants could also opt to have their computer monitored (e.g., screentime), but this data is not the focus of the current study, so it is not discussed.

Onboarding and ESM

In the experience sampling app, participants provided informed consent and received an intake questionnaire that measured their age, education status, and occupation, as well as other factors not of focal interest to the present study (a full method is available at our OSF page). Immediately after, participants began receiving ESM questionnaires according to a mixed-sampling scheme that contained semi-random and fixed elements. Following the initial notification, each ESM questionnaire remained available to the participant for 45 min. A reminder was sent out if the questionnaire had not been responded to 30 min after the initial notification. In total, participants received ESM questionnaires for 14 full days. After completing the ESM phase, participants received an outtake questionnaire that also measured factors that are not of direct interest to the present study (e.g., Trait Anxiety).

The first questionnaire of the day contained twenty-six items, the last questionnaire thirty-four, and the four in-between twenty-four items. Each questionnaire was identical in content, apart from the question's stem. Also, each questionnaire was presented to participants in a random order, to counter order effects.

Data preparation

Raw smartphone trace data was processed using Python packages pandas and Dask, and afterwards merged with the ESM data. This combined dataset enabled calculation of the three behavioral features. Additional data preparation (e.g., centering, lagging, etc.), as well as formal analyses, were conducted in R, using the packages dplyr, esmpack, misty (for grand-mean centering), and lavaan (for CFA and multilevel SEM models).

We first examined data for outliers and data errors, and whether model assumptions were met for later analyses (e.g., linearity). Then, select variables (e.g., mental fatigue) were lagged to the next dataset row to enable modeling across adjoining timepoints. For each participant, variables were only lagged when the next dataset row also represented the next received questionnaire for that person (thus, data from questionnaire 4, for instance, would only be able to predict data from questionnaire 5 for that same participant), and belonged to the same day (thus, data from the last questionnaire of the day could not predict data from the first questionnaire of the next day, an important analytic decision given the extensive temporal gap that would otherwise result between these questionnaires).

Following, level 2 versions of select level 1 variables (e.g., mental fatigue) were created to enable later analyses (i.e., the level 2 variable represented the average of level 1 variable values for a given participant). Level 1 (observation level) variables were then person-mean centered, to ensure model coefficients (in our Multilevel SEM) characterized the pure within-person relationship (i.e., with between-person variance removed). The centered behavioral features were additionally standardized at the within-person level to ensure model variables shared similar scales. Level 2 (person-level) variables were grand-mean centered to facilitate intercept interpretation. However, level 2 variables were only needed to saturate the level 2 model and enable multilevel SEM functionality with lavaan.

All code used in processing the data can be found on osf, under Data and Code / code. Please refer to the folder's README.txt for further information on which file contains which data processing steps.

The Android tracking app encountered issues on certain devices (Xiaomi, Huawei and Samsung). The main issue was the smartphone OS restricting access to different tracking data after a certain period of time. In response, we set up a monitoring dashboard to look for instances of potential data loss and could then troubleshoot with our participants. By providing device-specific instructions (see https://dontkillmyapp.com/) we could often pinpoint the cause and restore access.

Original Models

To keep the results section concise, only the results of the finalized models were discussed. The original models (with poor model fit), and the step-by-step procedure is detailed on osf. If you would like to run through the code itself, you can find an R Markdown file under Data and Code / code / 06_structural_models.Rmd. Alternatively, you can find a rendered version of this markdown file with the output of these models under: Data and Code / code / rendered_notebooks / 06_structural_models.html. This html page takes you through all the data analysis steps, informing about all analytical decisions taken.

Appendix B — Supplementary Material for Chapter 4

Mediation Models Output

Table 8.1: Full mediation model output.

Outcome	Predictor	B	SE	LL	UL	p	β
chat duration							
rushed	chat duration (c)	−.001	.007	−.014	.013	.916	−.000
rushed	work activity	.471	.045	.383	.558	.000	.138
rushed	morning	−.016	.019	−.054	.021	.398	−.004
rushed	evening	−.005	.019	−.041	.032	.795	−.001
juggling load	chat duration (a)	.025	.003	.019	.032	.000	.037
juggling load	work activity	.594	.016	.563	.625	.000	.431
juggling load	morning	−.033	.008	−.048	−.017	.000	−.022
juggling load	evening	.008	.008	−.009	.024	.362	.005
rushed	juggling load (b)	.839	.035	.771	.907	.000	.339
ab	a*b	.021	.003	.016	.027	.000	.013
total	c+(a*b)	.021	.007	.007	.034	.003	.012
chat frequency							
rushed	chat frequency (c)	−.003	.007	−.016	.011	.700	−.002
rushed	work activity	.471	.045	.383	.558	.000	.138
rushed	morning	−.017	.019	−.055	.021	.383	−.005
rushed	evening	−.005	.019	−.041	.032	.804	−.001
juggling load	chat frequency (a)	.023	.003	.016	.029	.000	.033
juggling load	work activity	.594	.016	.563	.625	.000	.431
juggling load	morning	−.031	.008	−.047	−.016	.000	−.021
juggling load	evening	.008	.008	−.008	.024	.336	.006
rushed	juggling load (b)	.839	.035	.771	.907	.000	.339
ab	a*b	.019	.003	.014	.024	.000	.011
total	c+(a*b)	.016	.007	.002	.030	.024	.010
chat notifications							
rushed	chat notifications (c)	.014	.007	−.000	.029	.053	.008
rushed	work activity	.468	.045	.381	.556	.000	.137
rushed	morning	−.011	.019	−.049	.027	.565	−.003
rushed	evening	−.006	.019	−.043	.030	.738	−.002
juggling load	chat notifications (a)	.028	.003	.022	.034	.000	.041
juggling load	work activity	.591	.016	.559	.622	.000	.428
juggling load	morning	−.028	.008	−.044	−.012	.000	−.019
juggling load	evening	.008	.008	−.008	.024	.336	.006
rushed	juggling load (b)	.838	.035	.770	.906	.000	.339
ab	a*b	.023	.003	.018	.029	.000	.014
total	c+(a*b)	.038	.008	.023	.053	.000	.022
chat fragmentation							
rushed	chat fragmentation (c)	.003	.006	−.010	.015	.686	.002

rushed	work activity	.470	.045	.383	.558	.000	.138
rushed	morning	−.016	.019	−.053	.022	.418	−.004
rushed	evening	−.005	.019	−.042	.031	.778	−.002
juggling load	chat fragmentation (a)	.026	.003	.019	.032	.000	.038
juggling load	work activity	.594	.016	.562	.625	.000	.430
juggling load	morning	−.033	.008	−.049	−.017	.000	−.022
juggling load	evening	.007	.008	−.009	.023	.387	.005
rushed	juggling load (b)	.839	.035	.770	.907	.000	.339
ab	a*b	.022	.003	.016	.027	.000	.013
total	c+(a*b)	.024	.007	.011	.037	.000	.014
social duration							
rushed	social duration (c)	.003	.007	−.011	.016	.713	.001
rushed	work activity	.471	.045	.383	.558	.000	.138
rushed	morning	−.016	.019	−.053	.022	.406	−.004
rushed	evening	−.005	.019	−.042	.031	.777	−.002
juggling load	social duration (a)	.007	.003	.001	.014	.024	.010
juggling load	work activity	.597	.016	.565	.628	.000	.433
juggling load	morning	−.038	.008	−.053	−.022	.000	−.026
juggling load	evening	.010	.008	−.007	.026	.244	.007
rushed	juggling load (b)	.839	.035	.771	.907	.000	.339
ab	a*b	.006	.003	.001	.011	.024	.003
total	c+(a*b)	.009	.007	−.006	.023	.245	.005
social frequency							
rushed	social frequency (c)	−.007	.007	−.021	.006	.295	−.004
rushed	work activity	.470	.045	.383	.558	.000	.138
rushed	morning	−.017	.019	−.055	.021	.375	−.005
rushed	evening	−.004	.019	−.040	.033	.838	−.001
juggling load	social frequency (a)	.009	.003	.003	.016	.005	.013
juggling load	work activity	.596	.016	.565	.628	.000	.432
juggling load	morning	−.037	.008	−.053	−.021	.000	−.025
juggling load	evening	.009	.008	−.007	.025	.271	.006
rushed	juggling load (b)	.839	.035	.771	.907	.000	.339
ab	a*b	.008	.003	.002	.013	.005	.004
total	c+(a*b)	.001	.007	−.014	.015	.937	.000
social notifications							
rushed	social notifications (c)	−.005	.008	−.021	.011	.545	−.002
rushed	work activity	.471	.045	.384	.558	.000	.138
rushed	morning	−.017	.019	−.054	.021	.382	−.005
rushed	evening	−.004	.019	−.041	.032	.816	−.001
juggling load	social notifications (a)	.006	.004	−.001	.014	.093	.007
juggling load	work activity	.596	.016	.565	.627	.000	.432
juggling load	morning	−.037	.008	−.053	−.021	.000	−.025
juggling load	evening	.010	.008	−.006	.026	.237	.007

rushed	juggling load (b)	.839	.035	.771	.907	.000	.339
ab	a*b	.005	.003	−.001	.012	.093	.002
total	c+(a*b)	.000	.009	−.017	.017	.971	.000
social fragmentation							
rushed	social fragmentation (c)	.004	.007	−.010	.017	.574	.002
rushed	work activity	.471	.045	.383	.558	.000	.138
rushed	morning	−.016	.019	−.053	.022	.412	−.004
rushed	evening	−.005	.019	−.042	.031	.770	−.002
juggling load	social fragmentation (a)	.011	.003	.005	.018	.001	.015
juggling load	work activity	.596	.016	.565	.628	.000	.432
juggling load	morning	−.037	.008	−.053	−.021	.000	−.025
juggling load	evening	.009	.008	−.007	.025	.271	.006
rushed	juggling load (b)	.839	.035	.771	.907	.000	.339
ab	a*b	.009	.003	.004	.015	.001	.005
total	c+(a*b)	.013	.007	−.001	.027	.068	.007
email duration							
rushed	email duration (c)	.018	.006	.006	.030	.004	.010
rushed	work activity	.469	.045	.382	.556	.000	.138
rushed	morning	−.014	.019	−.051	.024	.467	−.004
rushed	evening	−.006	.019	−.042	.031	.754	−.002
juggling load	email duration (a)	.016	.003	.010	.023	.000	.023
juggling load	work activity	.595	.016	.563	.626	.000	.431
juggling load	morning	−.036	.008	−.052	−.020	.000	−.025
juggling load	evening	.010	.008	−.007	.026	.239	.007
rushed	juggling load (b)	.838	.035	.770	.906	.000	.339
ab	a*b	.014	.003	.008	.019	.000	.008
total	c+(a*b)	.031	.007	.019	.044	.000	.018
email frequency							
rushed	email frequency (c)	.015	.006	.002	.028	.021	.009
rushed	work activity	.469	.045	.382	.557	.000	.138
rushed	morning	−.013	.019	−.051	.024	.486	−.004
rushed	evening	−.006	.019	−.042	.031	.758	−.002
juggling load	email frequency (a)	.019	.003	.013	.025	.000	.027
juggling load	work activity	.594	.016	.563	.625	.000	.431
juggling load	morning	−.034	.008	−.050	−.019	.000	−.024
juggling load	evening	.010	.008	−.007	.026	.249	.007
rushed	juggling load (b)	.838	.035	.770	.906	.000	.339
ab	a*b	.016	.003	.011	.021	.000	.009
total	c+(a*b)	.031	.007	.017	.045	.000	.018
email notifications							
rushed	email notifications (c)	.021	.009	.003	.039	.025	.010
rushed	work activity	.466	.045	.379	.554	.000	.137
rushed	morning	−.013	.019	−.051	.024	.488	−.004

rushed	evening	−.004	.019	−.041	.032	.827	−.001
juggling load	email notifications (a)	.030	.004	.021	.038	.000	.035
juggling load	work activity	.589	.016	.558	.621	.000	.427
juggling load	morning	−.034	.008	−.050	−.018	.000	−.023
juggling load	evening	.012	.008	−.004	.028	.155	.008
rushed	juggling load (b)	.838	.035	.770	.906	.000	.339
ab	a*b	.025	.004	.018	.032	.000	.012
total	c+(a*b)	.046	.010	.026	.065	.000	.022
email fragmentation							
rushed	email fragmentation (c)	.018	.006	.006	.030	.003	.010
rushed	work activity	.469	.045	.382	.556	.000	.137
rushed	morning	−.014	.019	−.051	.024	.480	−.004
rushed	evening	−.006	.019	−.043	.030	.741	−.002
juggling load	email fragmentation (a)	.018	.003	.011	.024	.000	.025
juggling load	work activity	.594	.016	.563	.625	.000	.431
juggling load	morning	−.035	.008	−.051	−.020	.000	−.024
juggling load	evening	.009	.008	−.007	.026	.258	.007
rushed	juggling load (b)	.838	.035	.770	.906	.000	.339
ab	a*b	.015	.003	.009	.020	.000	.009
total	c+(a*b)	.033	.006	.020	.045	.000	.019
w. comm. duration							
rushed	w. comm. duration (c)	.010	.012	−.013	.033	.379	.003
rushed	work activity	.470	.045	.383	.557	.000	.138
rushed	morning	−.016	.019	−.053	.022	.406	−.004
rushed	evening	−.005	.019	−.041	.032	.792	−.001
juggling load	w. comm. duration (a)	.022	.006	.010	.034	.000	.017
juggling load	work activity	.595	.016	.563	.626	.000	.431
juggling load	morning	−.038	.008	−.054	−.022	.000	−.026
juggling load	evening	.011	.008	−.006	.027	.201	.007
rushed	juggling load (b)	.839	.035	.771	.907	.000	.339
ab	a*b	.019	.005	.009	.029	.000	.006
total	c+(a*b)	.029	.012	.005	.053	.017	.009
w. comm. frequency							
rushed	w. comm. frequency (c)	.008	.012	−.017	.032	.530	.002
rushed	work activity	.470	.045	.383	.558	.000	.138
rushed	morning	−.016	.019	−.053	.022	.406	−.004
rushed	evening	−.005	.019	−.041	.032	.793	−.001
juggling load	w. comm. frequency (a)	.023	.006	.010	.035	.000	.017
juggling load	work activity	.594	.016	.563	.626	.000	.431
juggling load	morning	−.037	.008	−.053	−.022	.000	−.026
juggling load	evening	.011	.008	−.006	.027	.198	.008
rushed	juggling load (b)	.839	.035	.771	.907	.000	.339
ab	a*b	.019	.005	.009	.029	.000	.006

total	c+(a*b)	.027	.013	.001	.053	.044	.008
w. comm. notifications							
rushed	w. comm. notifications (c)	.038	.017	.005	.071	.026	.010
rushed	work activity	.467	.045	.380	.555	.000	.137
rushed	morning	−.015	.019	−.053	.022	.423	−.004
rushed	evening	−.004	.019	−.040	.033	.833	−.001
juggling load	w. comm. notifications (a)	.030	.008	.015	.046	.000	.020
juggling load	work activity	.593	.016	.562	.625	.000	.430
juggling load	morning	−.037	.008	−.053	−.022	.000	−.026
juggling load	evening	.011	.008	−.005	.028	.169	.008
rushed	juggling load (b)	.838	.035	.770	.906	.000	.339
ab	a*b	.025	.007	.012	.038	.000	.007
total	c+(a*b)	.063	.018	.028	.098	.000	.017
w. comm. fragmentation							
rushed	w. comm. fragmentation (c)	.010	.011	−.013	.032	.407	.003
rushed	work activity	.470	.045	.383	.557	.000	.138
rushed	morning	−.016	.019	−.053	.022	.405	−.004
rushed	evening	−.005	.019	−.041	.032	.791	−.001
juggling load	w. comm. fragmentation (a)	.024	.006	.012	.035	.000	.018
juggling load	work activity	.595	.016	.563	.626	.000	.431
juggling load	morning	−.038	.008	−.053	−.022	.000	−.026
juggling load	evening	.011	.008	−.006	.027	.204	.007
rushed	juggling load (b)	.839	.035	.771	.907	.000	.339
ab	a*b	.020	.005	.010	.030	.000	.006
total	c+(a*b)	.029	.012	.006	.053	.013	.009
Average							
rushed	juggling load	.839	.035	.770	.907	.000	.339
juggling load	work activity	.594	.016	.563	.626	.000	.431
rushed	work activity	.470	.045	.382	.557	.000	.138
rushed	morning	−.015	.019	−.053	.022	.431	−.004
rushed	evening	−.005	.019	−.042	.032	.788	−.001
juggling load	morning	−.035	.008	−.051	−.019	.000	−.024
juggling load	evening	.010	.008	−.007	.026	.257	.007

(16 models: four features across four app categories)

Person-Specific Output

Table 8.2: Person-specific effects: Feeling Rushed.

	Mean	SD	2.5%	97.5%	neg (sig.) %	neg (non-sig.) %	non-existent %	pos (non-sig.) %	pos (sig.) %
Chat									

chat duration	.03	.09	−.16	.21	0.26	14.99	67.70	17.05	0.00
chat frequency	.02	.10	−.17	.22	0.65	18.99	60.59	19.25	0.52
chat notifications	.05	.11	−.17	.27	0.65	27.65	47.67	22.61	1.42
chat fragmentation	.03	.07	−.11	.17	0.00	7.75	82.56	9.69	0.00
Social									
social duration	.01	.09	−.16	.18	0.13	9.17	81.40	9.04	0.26
social frequency	.00	.09	−.16	.17	0.39	9.69	79.07	10.72	0.13
social notifications	.01	.08	−.15	.16	0.26	7.49	85.01	7.24	0.00
social fragmentation	.02	.07	−.13	.16	0.13	5.94	87.21	6.59	0.13
Email									
email duration	.04	.06	−.09	.16	0.00	3.62	88.63	7.75	0.00
email frequency	.04	.08	−.13	.20	0.00	16.93	65.63	16.80	0.65
email notifications	.06	.14	−.21	.33	0.65	23.77	54.78	18.99	1.81
email fragmentation	.04	.05	−.06	.14	0.00	1.16	94.06	4.65	0.13
Work communication									
w. comm. duration	.03	.05	−.07	.14	0.00	1.03	96.25	2.71	0.00
w. comm. frequency	.03	.09	−.14	.21	0.00	3.36	92.76	3.88	0.00
w. comm. notifications	.08	.15	−.21	.37	0.39	6.98	86.05	5.68	0.90
w. comm. fragmentation	.03	.04	−.04	.11	0.00	0.00	98.06	1.94	0.00

(parameter estimates, heterogeneity interval, and random effects distribution)

Table 8.3: Person-specific effects: Juggling Load.

	Mean	SD	2.5%	97.5%	neg (sig.) %	neg (non-sig.) %	non-existent %	pos (non-sig.) %	pos (sig.) %
Chat									
chat duration	.03	.05	−.06	.12	0.13	4.91	84.11	8.79	2.07
chat frequency	.03	.05	−.07	.12	0.39	7.49	80.49	9.17	2.45

chat notifications	.03	.05	−.07	.13	0.39	10.98	75.71	10.98	1.94
chat fragmentation	.03	.05	−.07	.13	0.00	6.07	80.62	11.50	1.81
Social									
social duration	.01	.03	−.06	.08	0.26	2.33	93.28	3.62	0.52
social frequency	.01	.04	−.07	.09	0.13	3.10	90.05	6.20	0.52
social notifications	.01	.03	−.06	.08	0.00	1.55	94.44	3.62	0.39
social fragmentation	.01	.04	−.06	.09	0.26	2.07	92.89	3.75	1.03
Email									
email duration	.02	.05	−.08	.11	0.00	2.07	87.21	9.43	1.29
email frequency	.02	.05	−.08	.12	0.13	4.39	83.33	10.21	1.94
email notifications	.03	.06	−.09	.15	0.52	19.51	63.31	12.66	4.01
email fragmentation	.02	.05	−.08	.12	0.00	2.33	86.18	10.47	1.03
Work communication									
w. comm. duration	.02	.05	−.07	.11	0.00	0.13	98.97	0.90	0.00
w. comm. frequency	.02	.05	−.07	.12	0.00	0.78	96.51	2.58	0.13
w. comm. notifications	.03	.07	−.10	.16	0.52	4.91	88.89	4.52	1.16
w. comm. fragmentation	.03	.04	−.06	.11	0.00	0.00	99.22	0.78	0.00

(parameter estimates, heterogeneity interval, and random effects distribution)

Table 8.4: Cross-level interactions: Fixed effects estimates.

	Rushed — Chat Notifications				**Rushed — Email Notifications**			
	Model 1	*Model 2*	*Model 3*	*Model 4*	*Model 5*	*Model 6*	*Model 7*	*Model 8*
Intercept	3.504***	2.780***	2.897***	2.725***	3.504***	2.780***	2.897***	2.725***
Chat / Email	0.153***	0.090***	0.084***	0.044	0.145***	0.152***	0.112***	0.095**
Age	−0.015***				−0.015***			
Chat/Email × Age	−0.002*				−0.000			
Gender		0.271***				0.271***		
Chat/Email × Gender		−0.000				−0.023		
Parenthood			0.074				0.074	
Chat/Email × Parenthood			0.011				0.046*	
Segmentation				0.093*				0.093*
Chat/Email × Segmentation				0.020*				0.021
Participants	772	763	772	622	772	763	772	622
Observations	42,761	42,355	42,761	35,051	42,761	42,355	42,761	35,051
Marginal R^2 / Conditional R^2	0.012 / 0.418	0.009 / 0.419	0.003 / 0.418	0.006 / 0.393	0.013 / 0.420	0.011 / 0.421	0.005 / 0.420	0.008 / 0.395
	Juggling Load — Chat Notifications				**Juggling Load — Email Notifications**			
	Model 9	*Model 10*	*Model 11*	*Model 12*	*Model 13*	*Model 14*	*Model 15*	*Model 16*
Intercept	0.559***	0.493***	0.417***	0.426***	0.559***	0.492***	0.416***	0.427***
Chat / Email	0.097***	0.065***	0.058***	0.033**	0.109***	0.094***	0.081***	0.041*
Age	−0.001				−0.001			
Chat/Email × Age	−0.001*				−0.001			
Gender		0.032				0.033		
Chat/Email × Gender		−0.004				−0.005		
Parenthood			0.167***				0.167***	
Chat/Email × Parenthood			0.008				0.016	
Segmentation				0.044*				0.044***
Chat/Email × Segmentation				0.012*				0.022***

Participants	772	763	772	622	772	763	772	622
Observations	42,764	42,358	42,764	35,053	42,764	42,358	42,764	35,053
Marginal R^2 / Conditional R^2	0.009 / 0.208	0.009 / 0.207	0.023 / 0.207	0.012 / 0.187	0.012 / 0.212	0.012 / 0.212	0.026 / 0.212	0.016 / 0.193

* $p < .05$ ** $p < .01$ *** $p < .001$. Cross-level interactions used *unstandardized* dependent variables (rushed, juggling load) and *within-person standardized* smartphone measures (chat and email notifications). Our OSF page includes the full model output.

Appendix C — Supplementary Material for Chapter 6

Step-by-Step Guide for Collecting Log Data from iOS, macOS, and iPadOS

Preparation

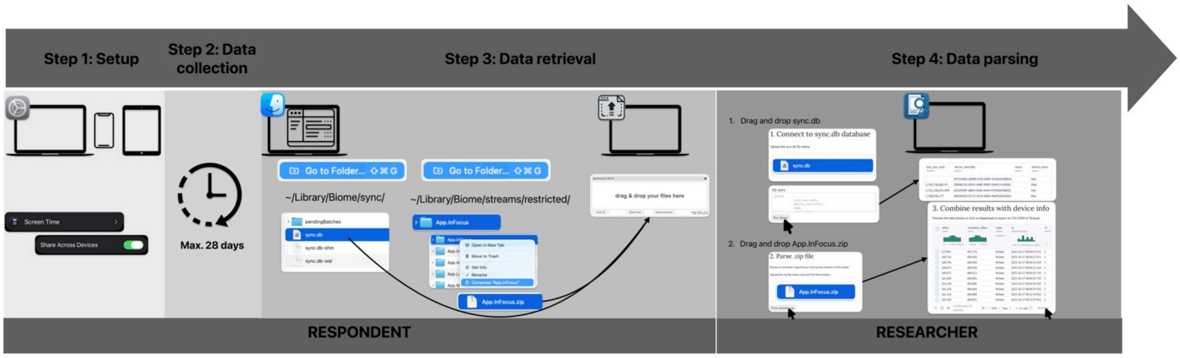
Device Requirements:

- iOS (minimum iOS 26) and iPadOS (minimum iPadOS 26) data can only be captured if the respondent also has a Mac (minimum MacOS 26) linked to the same Apple ID. The device of the Researcher can be a Mac or Windows.

Tool Requirements:

- You'll need an internet connection to access the marimo notebook OR install docker and download the tool to parse the files locally.

To use the tool:



- Navigate to the publicly available Marimo Notebook: https://molab.marimo.io/notebooks/nb_iVkKovhPYZBGnMVggpXWSL/app
  - Drag and drop the collected files of the participants

OR

- Run a local instance of the Marimo Notebook through docker:
  - Install Docker desktop
  - Download the tool from the Project OSF: https://osf.io/6puxm/?view_only=b44063a220204bdca33b964e2e8a4ebd
  - Run from the Root Repo:

```
docker load -i ios-biome-screen-time.tar
docker run -p 8080:8080 kvgaever/ios-biome-screen-time:latest
```

- Go to http://localhost:8080
- Drag and drop the collected files of the participants

Step 1. Set-up — On Device

Ensure all devices use the same Apple ID. Enable Screen Time Synchronization:

On Mac:

1. Open **System Preferences**.
2. Click on **Screen Time**.
3. Scroll down and toggle **Share Across Devices** to **ON**.

On iPhone:

1. Open **Settings**.
2. Tap on **Screen Time**.
3. Scroll down and toggle **Share Across Devices** to **ON**.

On iPad:

1. Open **Settings**.
2. Tap on **Screen Time**.
3. Scroll down and toggle **Share Across Devices** to **ON**.

On Apple Watch: No extra settings need to be changed.

Once synchronization is enabled, the Mac will store a file containing log data from all the devices that are synchronized.

Step 2. Data Collection

Wait for the desired logging period (e.g., 14 days) to collect comprehensive data. The database will store use data of the last 4 weeks or since synchronization was enabled, whichever is shorter.

Step 3. Data Retrieval — On Device and Transfer

Locate Sync.db:

1. Open **Finder** on the Mac.
2. From the top menu, select **Go > Go to Folder...**
3. Enter the following path: `~/Library/Biome/sync/Sync.db`

Locate and compress the biome files:

1. Open **Finder** on the Mac.
2. From the top menu, select **Go > Go to Folder...**
3. Enter the following path: `~/Library/Biome/streams/restricted/`
4. Locate the folder named **App.InFocus**.
5. Zip the App.InFocus folder.

Retrieve the folder and the Sync.db file: Let the respondents donate the compressed folder and the Sync.db file using a data donation method.

Step 4. Data Parsing — On Device of Researcher

Open the Tool & drag and drop the sync.db file and the compressed folder to the right location and download the dataset in a usable format.

References

Aalbers, G. (2023). *Digital traces of the mind: Using smartphones to capture signals of well-being in individuals*. Tilburg University.

Aalbers, G., Hendrickson, A. T., Vanden Abeele, M. M. P., & Keijsers, L. (2023). Smartphone-Tracked Digital Markers of Momentary Subjective Stress in College Students: Idiographic Machine Learning Analysis. *JMIR mHealth and uHealth*, *11*(1), e37469. https://doi.org/10.2196/37469

Aalbers, G., Vanden Abeele, M. M. P., Hendrickson, A. T., De Marez, L., & Keijsers, L. (2021). Caught in the moment: Are there person-specific associations between momentary procrastination and passively measured smartphone use? *Mobile Media & Communication*, 2050157921993896. https://doi.org/10.1177/2050157921993896

Abrahamse, E., Parry, D. A., le Roux, D. B., & Baumgartner, S. (2025, August 12). *Media-Multitasking And Attentional Control: A Theory-Driven Research Agenda* (nj4sz_v1). https://doi.org/10.31234/osf.io/nj4sz_v1 (Pre-published)

Alexander, B., Ettema, D., & Dijst, M. (2010). Fragmentation of work activity as a multi-dimensional construct and its association with ICT, employment and sociodemographic characteristics. *Journal of Transport Geography*, *18*(1), 55–64. https://doi.org/10.1016/j.jtrangeo.2009.05.010

Alexander, B., Hubers, C., Schwanen, T., Dijst, M., & Ettema, D. (2011). Anything, anywhere, anytime? Developing indicators to assess the spatial and temporal fragmentation of activities. *Environment and Planning B: Planning and Design*, *38*(4), 678–705. https://doi.org/10.1068/b35132

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), 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

Ashforth, B. E., Kreiner, G. E., & Fugate, M. (2000). All in a Day's Work: Boundaries and Micro Role Transitions. *The Academy of Management Review*, *25*(3), 472–491. https://doi.org/10.2307/259305

Baethge, A., & and Rigotti, T. (2013). Interruptions to workflow: Their relationship with irritation and satisfaction with performance, and the mediating roles of time pressure and mental demands. *Work & Stress*, *27*(1), 43–63. https://doi.org/10.1080/02678373.2013.761783

Balliu, B., Douglas, C., Seok, D., Shenhav, L., Wu, Y., Chatzopoulou, D., Kaiser, W., Chen, V., Kim, J., Deverasetty, S., Arnaudova, I., Gibbons, R., Congdon, E., Craske, M. G., Freimer, N., Halperin, E., Sankararaman, S., & Flint, J. (2024). Personalized mood prediction from patterns of behavior collected with smartphones. *Npj Digital Medicine*, *7*(1), 49. https://doi.org/10.1038/s41746-024-01035-6

Bates, D., Mächler, M., Bolker, B., & Walker, S. (2015). Fitting linear mixed-effects models using lme4. *Journal of Statistical Software*, *67*(1), 1–48. https://doi.org/10.18637/jss.v067.i01

Baumgartner, S. E. (2025). On the stabilization of media effects after repeated exposure: Consequences for media effects research. *Communication Theory*, qtaf017. https://doi.org/10.1093/ct/qtaf017

Baumgartner, S. E., & Sumter, S. R. (2017). Dealing with media distractions: An observational study of computer-based multitasking among children and adults in the Netherlands. *Journal of Children and Media*, *11*(3), 295–313. https://doi.org/10.1080/17482798.2017.1304971

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

Bayer, J. B., Campbell, S. W., & Ling, R. (2016). Connection Cues: Activating the Norms and Habits of Social Connectedness. *Communication Theory*, *26*(2), 128–149. https://doi.org/10.1111/comt.12090

Bayer, J. B., Triệu, P., & Ellison, N. B. (2020). Social Media Elements, Ecologies, and Effects. *Annual Review of Psychology*, *71*, 471–497. https://doi.org/10.1146/annurev-psych-010419-050944

Belkin, L. Y., Becker, W. J., & Conroy, S. A. (2020). The Invisible Leash: The Impact of Organizational Expectations for Email Monitoring After-Hours on Employee Resources, Well-Being, and Turnover Intentions. *Group & Organization Management*, *45*(5), 709–740. https://doi.org/10.1177/1059601120933143

Berry, D. M. (2011). The Computational Turn: Thinking About the Digital Humanities. *Culture Machine*, *12*.

Bittman, M., Brown, J. E., & Wajcman, J. (2009). The mobile phone, perpetual contact and time pressure. *Work, Employment and Society*, *23*(4), 673–691. https://doi.org/10.1177/0950017009344910

Bjerre-Nielsen, A., & Glavind, K. L. (2022). Ethnographic data in the age of big data: How to compare and combine. *Big Data & Society*, *9*(1), 20539517211069893. https://doi.org/10.1177/20539517211069893

Bleier, A. (2025). *What is Computational Reproducibility? (GESIS Guides to Digital Behavioral Data, 2)* (Version 1.0). GESIS - Leibniz Institute for the Social Sciences. https://doi.org/10.60762/GGDBD25002.1.0

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

Bosch, O. J., Sturgis, P., Kuha, J., & Revilla, M. (2025). Uncovering Digital Trace Data Biases: Tracking Undercoverage in Web Tracking Data. *Communication Methods and Measures*, *19*(2), 157–177. https://doi.org/10.1080/19312458.2024.2393165

Bossio, D., & Holton, A. E. (2021). Burning out and turning off: Journalists' disconnection strategies on social media. *Journalism*, *22*(10), 2475–2492. https://doi.org/10.1177/1464884919872076

boyd, danah, & Crawford, K. (2012). Critical Questions for Big Data. *Information, Communication & Society*, *15*(5), 662–679. https://doi.org/10.1080/1369118X.2012.678878

Boyle, S. C., Baez, S., Trager, B. M., LaBrie, J. W., Boyle, S. C., Baez, S., Trager, B. M., & LaBrie, J. W. (2022). Systematic Bias in Self-Reported Social Media Use in the Age of Platform Swinging: Implications for Studying Social Media Use in Relation to Adolescent Health Behavior. *International Journal of Environmental Research and Public Health*, *19*(16). https://doi.org/10.3390/ijerph19169847

Brinberg, M., Ram, N., Wang, J., Sundar, S. S., Cummings, J. J., Yeykelis, L., & Reeves, B. (2023). Screenertia: Understanding "Stickiness" of Media Through Temporal Changes in Screen Use. *Communication Research*, *50*(5), 535–560. https://doi.org/10.1177/00936502211062778

Bröhl, C., Rasche, P., Jablonski, J., Theis, S., Wille, M., & Mertens, A. (2018). Desktop PC, Tablet PC, or Smartphone? An Analysis of Use Preferences in Daily Activities for Different Technology Generations of a Worldwide Sample. In J. Zhou & G. Salvendy (Eds.), *Human Aspects of IT for the Aged Population. Acceptance, Communication and Participation* (Vol. 10926, pp. 3–20). Springer International Publishing. https://doi.org/10.1007/978-3-319-92034-4_1

Buda, T. S., Khwaja, M., Garriga, R., & Matic, A. (2023). Two edges of the screen: Unpacking positive and negative associations between phone use in everyday contexts and subjective well-being. *PLOS ONE*, *18*(4), e0284104. https://doi.org/10.1371/journal.pone.0284104

Castells, M. (Ed.). (2007). *Mobile communication and society: A global perspective: A project of the Annenberg Research Network on international communication*. MIT Press.

Chan, M. (2015). Multimodal Connectedness and Quality of Life: Examining the Influences of Technology Adoption and Interpersonal Communication on Well-Being Across the Life Span. *Journal of Computer-Mediated Communication*, *20*(1), 3–18. https://doi.org/10.1111/jcc4.12089

Chan, M., & Yi, J. (2024). Social Media Use and Political Engagement in Polarized Times. Examining the Contextual Roles of Issue and Affective Polarization in Developed Democracies. *Political Communication*, *41*(5), 743–762. https://doi.org/10.1080/10584609.2024.2325423

Conrad, F. G., Keusch, F., & Schober, M. F. (2021). New Data in Social and Behavioral Research. *Public Opinion Quarterly*, *85*(S1), 253–263. https://doi.org/10.1093/poq/nfab027

Cornwell, B. (2013). Switching Dynamics and the Stress Process. *Social Psychology Quarterly*, *76*(2), 99–124. https://doi.org/10.1177/0190272513482133

Courage, M. L., Bakhtiar, A., Fitzpatrick, C., Kenny, S., & Brandeau, K. (2015). Growing up multitasking: The costs and benefits for cognitive development. *Developmental Review, Special Issue: Living in the "Net" Generation: Multitasking, Learning, and Development*, *35*, 5–41. https://doi.org/10.1016/j.dr.2014.12.002

De Hert, S. (2020). Burnout in Healthcare Workers: Prevalence, Impact and Preventative Strategies. *Local and Regional Anesthesia*, *13*, 171–183. https://doi.org/10.2147/LRA.S240564

De Marez, L., Sevenhant, R., Denecker, F., Georges, A., Wuyts, G., & Schuurman, D. (2022). *Imec.digimeter 2022* [Digimeter]. imec. https://www.imec.be/sites/default/files/2023-03/imec_digimeter_2022.pdf de Segovia Vicente, D., Van Gaeveren, K., Murphy, S. L., & Vanden Abeele, M. M. P. (2024). Does mindless scrolling hamper well-being? Combining ESM and log-data to examine the link between mindless scrolling, goal conflict, guilt, and daily well-being. *Journal of Computer-Mediated Communication*, *29*(1), zmad056. https://doi.org/10.1093/jcmc/zmad056

Demerouti, E., Bakker, A., Nachreiner, F., & Ebbinghaus, M. (2002). From mental strain to burnout. *European Journal of Work and Organizational Psychology*, *11*(4), 423–441. https://doi.org/10.1080/13594320244000274

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

Denovan, A., Dagnall, N., & Drinkwater, K. (2024). The Chronic Time Pressure Inventory: Further assessment of factorial structure and validity. *PeerJ*, *12*, e17373. https://doi.org/10.7717/peerj.17373

Derks, D., Bakker, A. B., & Gorgievski, M. (2021). Private smartphone use during worktime: A diary study on the unexplored costs of integrating the work and family domains. *Computers in Human Behavior*, *114*. https://doi.org/10.1016/j.chb.2020.106530

Duxbury, L., Higgins, C., Smart, R., & Stevenson, M. (2014). Mobile Technology and Boundary Permeability. *British Journal of Management*, *25*(3), 570–588. https://doi.org/10.1111/1467-8551.12027

Duxbury, L., Stevenson, M., & Higgins, C. (2017). Too much to do, too little time: Role overload and stress in a multi-role environment. *International Journal of Stress Management*, *25*(3), 250–266. https://doi.org/10.1037/str0000062

Edú-Valsania, S., Laguía, A., & Moriano, J. A. (2022). Burnout: A Review of Theory and Measurement. *International Journal of Environmental Research and Public Health*, *19*(3), 1780. https://doi.org/10.3390/ijerph19031780

Eisert, J., Di Nocera, F., Baldwin, C., Lee, J., Stephen Higgins, J., Helton, W. S., & Hancock, P. (2016). Vigilance and Fatigue: A Double Sided Coin? *Proceedings of the Human Factors and Ergonomics Society Annual Meeting*, *60*(1), 1563–1568. https://doi.org/10.1177/1541931213601361

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

Elmer, T., Fernández, A., Stadel, M., Kas, M. J. H., & Langener, A. M. (2025). Bidirectional associations between smartphone usage and momentary well-being in young adults: Tackling methodological challenges by combining experience sampling methods with passive smartphone data. *Emotion*. https://doi.org/10.1037/emo0001485

Enders, C. K., & Tofighi, D. (2007). Centering predictor variables in cross-sectional multilevel models: A new look at an old issue. *Psychological Methods*, *12*(2), 121–138. https://doi.org/10.1037/1082-989X.12.2.121

Ernala, S. K., Burke, M., Leavitt, A., & Ellison, N. B. (2022). Mindsets Matter: How Beliefs About Facebook Moderate the Association Between Time Spent and Well-Being. *CHI Conference on Human Factors in Computing Systems*, 1–13. https://doi.org/10.1145/3491102.3517569

Es, K. V., Schäfer, M. T., & Wieringa, M. (2021). Tool Criticism and the Computational Turn. A "Methodological Moment" in Media and Communication Studies. *Medien & Kommunikationswissenschaft*, *69*(1), 46–64. https://doi.org/10.5771/1615-634X-2021-1-46

Fast, K., Lindell, J., & Jansson, A. (2021). Disconnection as Distinction: A Bourdieusian Study of Where People Withdraw from Digital Media. In A. Jansson & P. C. Adams (Eds.), *Disentangling: The Geographies of Digital Disconnection* (pp. 61–90). Oxford University Press. https://doi.org/10.1093/oso/9780197571873.003.0004

Ferguson, C. J. (2025). Smartphone bans in schools remain unproven. *World Journal of Pediatrics*, *21*(8), 769–774. https://doi.org/10.1007/s12519-025-00951-1

Ferguson, C. J. (2026). Links between social media use and mental wellness in youth are an artifact of other factors: Implications for public policy and meta-analysis. *Current Psychology*, *45*(6), 573. https://doi.org/10.1007/s12144-026-09205-3

Ferguson, C. J., Kaye, L. K., Branley-Bell, D., Markey, P., Ivory, J. D., Klisanin, D., Elson, M., Smyth, M., Hogg, J. L., McDonnell, D., Nichols, D., Siddiqui, S., Gregerson, M., & Wilson, J. (2022). Like this meta-analysis: Screen media and mental health. *Professional Psychology: Research and Practice*, *53*, 205–214. https://doi.org/10.1037/pro0000426

Finomore, V. S., Warm, J. S., Matthews, G., Riley, M. A., Dember, W. N., Shaw, T. H., Ungar, N. R., & Scerbo, M. W. (2006). Measuring the Workload of Sustained Attention. *Proceedings of the Human Factors and Ergonomics Society Annual Meeting*, *50*(16), 1614–1618. https://doi.org/10.1177/154193120605001621

Freytag, A., Knop-Huelss, K., Meier, A., Reinecke, L., Hefner, D., Klimmt, C., & Vorderer, P. (2021). Permanently Online—Always Stressed Out? The Effects of Permanent Connectedness on Stress Experiences. *Human Communication Research*, *47*(2), 132–165. https://doi.org/10.1093/hcr/hqaa014

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

Gilbert, A., Baumgartner, S. E., & Reinecke, L. (2023). Situational boundary conditions of digital stress: Goal conflict and autonomy frustration make smartphone use more stressful. *Mobile Media & Communication*, *11*(3), 435–458. https://doi.org/10.1177/20501579221138017

Goedhart, G., Kromhout, H., Wiart, J., & Vermeulen, R. (2015). Validating self-reported mobile phone use in adults using a newly developed smartphone application. *Occupational and Environmental Medicine*, *72*(11), 812–818. https://doi.org/10.1136/oemed-2015-102808

Goodyear, V. A., Randhawa, A., Adab, P., Al-Janabi, H., Fenton, S., Jones, K., Michail, M., Morrison, B., Patterson, P., Quinlan, J., Sitch, A., Twardochleb, R., Wade, M., & Pallan, M. (2025). School phone policies and their association with mental wellbeing, phone use, and social media use (SMART Schools): A cross-sectional observational study. *The Lancet Regional Health - Europe*, *51*, 101211. https://doi.org/10.1016/j.lanepe.2025.101211

Görland, S. O. (2019). Really "Dead Time"? Mobile Media Use and Time Perception in In-between Times. In M. Hartmann, E. Prommer, K. Deckner, & S. O. Görland (Eds.), *Mediated Time: Perspectives on Time in a Digital Age* (pp. 321–340). Springer International Publishing. https://doi.org/10.1007/978-3-030-24950-2_16

Grier, R. A., Warm, J. S., Dember, W. N., Matthews, G., Galinsky, T. L., Szalma, J. L., & Parasuraman, R. (2003). The Vigilance Decrement Reflects Limitations in Effortful Attention, Not Mindlessness. *Human Factors*, *45*(3), 349–359. https://doi.org/10.1518/hfes.45.3.349.27253

Griffioen, N., Rooij, M. van, Lichtwarck-Aschoff, A., & Granic, I. (2020). Toward Improved Methods in Social Media Research. *Technology, Mind, and Behavior*, *1*(1). https://doi.org/10.1037/tmb0000005

große Deters, F., & Schoedel, R. (2024). Keep on scrolling? Using intensive longitudinal smartphone sensing data to assess how everyday smartphone usage behaviors are related to well-being. *Computers in Human Behavior*, *150*, 107977. https://doi.org/10.1016/j.chb.2023.107977

Grotto, A. R., & Mills, M. J. (2023). Crossing the line: The violating effects of illegitimate interruptions from work and the differential impact on work-to-family conflict by gender. *Journal of Organizational Behavior*, *44*(4), 700–716. https://doi.org/10.1002/job.2689

H. Birk, R., & Samuel, G. (2020). Can digital data diagnose mental health problems? A sociological exploration of "digital phenotyping." *Sociology of Health & Illness*, *42*(8), 1873–1887. https://doi.org/10.1111/1467-9566.13175

Halfmann, A., Meier, A., & Reinecke, L. (2021). Too Much or Too Little Messaging? Situational Determinants of Guilt About Mobile Messaging. *Journal of Computer-Mediated Communication*, *26*(2), 72–90. https://doi.org/10.1093/jcmc/zmaa018

Halfmann, A., & Rieger, D. (2019). Permanently on Call: The Effects of Social Pressure on Smartphone Users' Self-Control, Need Satisfaction, and Well-Being. *Journal of Computer-Mediated Communication*, *24*(4), 165–181. https://doi.org/10.1093/jcmc/zmz008

Hall, J. (2017). The Experience of Mobile Entrapment in Daily Life. *Journal of Media Psychology*, *29*(3), 148–158. https://doi.org/10.1027/1864-1105/a000228

Hall, J. A. (2023). What we do in the shadows: The consumption of mobile messaging by social media mobile apps in the twilight of the social networking era. *Mobile Media & Communication*, *11*(1), 66–73. https://doi.org/10.1177/20501579221133610

Harari, G. M., & Gosling, S. D. (2023). Understanding behaviours in context using mobile sensing. *Nature Reviews. Psychology*, *2*(12), 767–779. https://doi.org/10.1038/s44159-023-00235-3

Harari, G. M., Müller, S. R., Aung, M. S., & Rentfrow, P. J. (2017). Smartphone sensing methods for studying behavior in everyday life. *Current Opinion in Behavioral Sciences, Big Data in the Behavioural Sciences*, *18*, 83–90. https://doi.org/10.1016/j.cobeha.2017.07.018

Haythornthwaite, C. (2005). Social networks and Internet connectivity effects. *Information, Communication & Society*, *8*(2), 125–147. https://doi.org/10.1080/13691180500146185

Head, J., & Helton, W. S. (2014). Sustained attention failures are primarily due to sustained cognitive load not task monotony. *Acta Psychologica*, *153*, 87–94. https://doi.org/10.1016/j.actpsy.2014.09.007

Hendrikx, K., Van Ruysseveldt, J., Proost, K., & van der Lee, S. (2023). "Out of office": Availability norms and feeling burned out during the COVID-19 pandemic: The mediating role of autonomy and telepressure. *Frontiers in Psychology*, *14*. https://www.frontiersin.org/articles/10.3389/fpsyg.2023.1063020

Hepp, A., Loosen, W., & Hasebrink, U. (2021). Jenseits des Computational Turn: Methodenentwicklung und Forschungssoftware in der Kommunikations- und Medienwissenschaft – zur Einführung in das Themenheft. *Medien & Kommunikationswissenschaft*, *69*(1), 3–24. https://doi.org/10.5771/1615-634X-2021-1-3-1

Hodes, L. N., & Thomas, K. G. F. (2021). Smartphone Screen Time: Inaccuracy of self-reports and influence of psychological and contextual factors. *Computers in Human Behavior*, *115*, 106616. https://doi.org/10.1016/j.chb.2020.106616

Hodgetts, H. M., Vachon, F., Chamberland, C., & Tremblay, S. (2017). See no evil: Cognitive challenges of security surveillance and monitoring. *Journal of Applied Research in Memory and Cognition*, *6*(3), 230–243. https://doi.org/10.1016/j.jarmac.2017.05.001

Hopstaken, J. F., van der Linden, D., Bakker, A. B., & Kompier, M. A. J. (2015). A multifaceted investigation of the link between mental fatigue and task disengagement. *Psychophysiology*, *52*(3), 305–315. https://doi.org/10.1111/psyp.12339

Hox, J. (1998). Multilevel Modeling: When and Why. In I. Balderjahn, R. Mathar, & M. Schader (Eds.), *Classification, Data Analysis, and Data Highways* (pp. 147–154). Springer. https://doi.org/10.1007/978-3-642-72087-1_17

Hu, L., & Bentler, P. M. (1999). Cutoff criteria for fit indexes in covariance structure analysis: Conventional criteria versus new alternatives. *Structural Equation Modeling: A Multidisciplinary Journal*, *6*(1), 1–55. https://doi.org/10.1080/10705519909540118

Inzlicht, M., Schmeichel, B. J., & Macrae, C. N. (2014). Why self-control seems (but may not be) limited. *Trends in Cognitive Sciences*, *18*(3), 127–133. https://doi.org/10.1016/j.tics.2013.12.009

Jeong, Y., Jung, H., & Lee, J. (2020). Cyberslacking or Smart Work: Smartphone Usage Log-Analysis Focused on App-Switching Behavior in Work and Leisure Conditions. *International Journal of Human-Computer Interaction*, *36*(1), 15–30. https://doi.org/10.1080/10447318.2019.1597574

Jo, H.-H., Karsai, M., Kertész, J., & Kaski, K. (2012). Circadian pattern and burstiness in mobile phone communication. *New Journal of Physics*, *14*(1), 013055. https://doi.org/10.1088/1367-2630/14/1/013055

Johannes, N. (2020). *Effects of Smartphone Cues and Online Vigilance on Well-Being and Performance* [PhD thesis, Radboud University]. https://repository.ubn.ru.nl/handle/2066/214893

Johannes, N., Meier, A., Reinecke, L., Ehlert, S., Setiawan, D. N., Walasek, N., Dienlin, T., Buijzen, M., & Veling, H. (2021). The relationship between online vigilance and affective well-being in everyday life: Combining smartphone logging with experience sampling. *Media Psychology*, *24*(5), 581–605. https://doi.org/10.1080/15213269.2020.1768122

Johannes, N., Nguyen, T., Weinstein, N., & Przybylski, A. K. (2021). Objective, Subjective, and Accurate Reporting of Social Media Use: No Evidence That Daily Social Media Use Correlates With Personality Traits, Motivational States, or Well-Being. *Technology, Mind, and Behavior*, *2*(2). https://doi.org/10.1037/tmb0000035

Johannes, N., Veling, H., Dora, J., Meier, A., Reinecke, L., & Buijzen, M. (2018). Mind-Wandering and Mindfulness as Mediators of the Relationship Between Online Vigilance and Well-Being. *Cyberpsychology, Behavior, and Social Networking*, *21*(12), 761–767. https://doi.org/10.1089/cyber.2018.0373

Johannes, N., Veling, H., Verwijmeren, T., & Buijzen, M. (2019). Hard to Resist? *Journal of Media Psychology*, *31*(4), 214–225. https://doi.org/10.1027/1864-1105/a000248

Kalman, Y. M., Ballard, D. I., & Aguilar, A. M. (2021). Chronemic urgency in everyday digital communication. *Time & Society*, *30*(2), 153–175. https://doi.org/10.1177/0961463X20987721

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

Keusch, F., & Conrad, F. G. (2022). Using Smartphones to Capture and Combine Self-Reports and Passively Measured Behavior in Social Research. *Journal of Survey Statistics and Methodology*, *10*(4), 863–885. https://doi.org/10.1093/jssam/smab035

Keusch, F., Pankowska, P. K., Cernat, A., & Bach, R. L. (2024). Do You Have Two Minutes to Talk about Your Data? Willingness to Participate and Nonparticipation Bias in Facebook Data Donation. *Field Methods*, *36*(4), 279–293. https://doi.org/10.1177/1525822X231225907

Keusch, F., Wenz, A., & Conrad, F. (2022). Do you have your smartphone with you? Behavioral barriers for measuring everyday activities with smartphone sensors. *Computers in Human Behavior*, *127*, 107054. https://doi.org/10.1016/j.chb.2021.107054

Klingelhoefer, J., Gilbert, A., Adrian, C., & Meier, A. (2025). Possible futures all at once: Time frame and time lag in short-term longitudinal media effects research on well-being. *Journal of Communication*, jqaf037. https://doi.org/10.1093/joc/jqaf037

Klingelhoefer, J., Gilbert, A., & Meier, A. (2023). *Momentary Motivations for Digital Disconnection: An Experience Sampling Study*. https://doi.org/10.31234/osf.io/qhctd

Klingelhoefer, J., Gilbert, A., & Meier, A. (2024). Momentary motivations for digital disconnection: An experience sampling study. *Journal of Computer-Mediated Communication*, *29*(5), zmae013. https://doi.org/10.1093/jcmc/zmae013

Lanctot, A., & Duxbury, L. (2021). When everything is urgent! Mail use and employee well-being. *Computers in Human Behavior Reports*, *4*, 100152. https://doi.org/10.1016/j.chbr.2021.100152

Lee, A. Y., & Hancock, J. T. (2024). Social media mindsets: A new approach to understanding social media use and psychological well-being. *Journal of Computer-Mediated Communication*, *29*(1), zmad048. https://doi.org/10.1093/jcmc/zmad048

Lee, A. Y., Katz, R., & Hancock, J. (2021). The Role of Subjective Construals on Reporting and Reasoning about Social Media Use. *Social Media + Society*, *7*(3), 20563051211035350. https://doi.org/10.1177/20563051211035350

Leitgöb, H., & Keusch, F. (2026). Causal Inferences from Digital Behavioral Data. *KZfSS Kölner Zeitschrift Für Soziologie Und Sozialpsychologie*. https://doi.org/10.1007/s11577-026-01050-3

Li, J., Li, H., Wang, H., Umer, W., Fu, H., & Xing, X. (2019). Evaluating the impact of mental fatigue on construction equipment operators' ability to detect hazards using wearable eye-tracking technology. *Automation in Construction*, *105*, 102835. https://doi.org/10.1016/j.autcon.2019.102835

Limniou, M. (2021). The Effect of Digital Device Usage on Student Academic Performance: A Case Study. *Education Sciences*, *11*(3), 121. https://doi.org/10.3390/educsci11030121

Ling, R. (2017). A Brief History of Individual Addressability: The Role of Mobile Communication in Being Permanently Connected. In *Permanently Online, Permanently Connected*. Routledge.

Liu, D., Baumeister, R. F., & Yang, C.-C. (2024). A meta-analysis on the relationship between the use of electronic media and psychological well-being. *Emerging Trends in Drugs, Addictions, and Health*, *4*, 100162. https://doi.org/10.1016/j.etdah.2024.100162

Lu, Z. (2024). Can Work Time Fragmentation Influence Workers' Subjective Time Pressure? The Roles of Gender and Parenthood. *Sociology*, *58*(1), 194–212. https://doi.org/10.1177/00380385231166893

Lutz, S., Schneider, F. M., & Halfmann, A. (2026). Staying healthy despite social pressure to be available? An experimental application of the Integrative Model of Mobile Media Use and Need Experiences (IM3UNE). *Computers in Human Behavior Reports*, *21*, 100977. https://doi.org/10.1016/j.chbr.2026.100977

MacCallum, R. C., Browne, M. W., & Sugawara, H. M. (1996). Power analysis and determination of sample size for covariance structure modeling. *Psychological Methods*, *1*(2), 130–149. https://doi.org/10.1037/1082-989X.1.2.130

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*, 107567. https://doi.org/10.1016/j.chb.2022.107567

Marciano, L., Driver, C. C., Schulz, P. J., & Camerini, A.-L. (2022). Dynamics of adolescents' smartphone use and well-being are positive but ephemeral. *Scientific Reports*, *12*(1), 1316. https://doi.org/10.1038/s41598-022-05291-y

Marciano, L., & Viswanath, K. (2023). Social media use and adolescents' well-being: A note on flourishing. *Frontiers in Psychology*, *14*. https://doi.org/10.3389/fpsyg.2023.1092109

Mark, G., Gudith, D., & Klocke, U. (2008). The cost of interrupted work: More speed and stress. *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems, CHI '08*, 107–110. https://doi.org/10.1145/1357054.1357072

Marsh, H. W., Wen, Z., & Hau, K.-T. (2004). Structural Equation Models of Latent Interactions: Evaluation of Alternative Estimation Strategies and Indicator Construction. *Psychological Methods*, *9*(3), 275–300. https://doi.org/10.1037/1082-989X.9.3.275

Matthews, G., Warm, J. S., Reinerman-Jones, L. E., Langheim, L. K., Washburn, D. A., & Tripp, L. (2010). Task engagement, cerebral blood flow velocity, and diagnostic monitoring for sustained attention. *Journal of Experimental Psychology: Applied*, *16*(2), 187–203. https://doi.org/10.1037/a0019572

McDowall, A., & Kinman, G. (2017). The new nowhere land? A research and practice agenda for the "always on" culture. *Journal of Organizational Effectiveness: People and Performance*, *4*(3), 256–266. https://doi.org/10.1108/JOEPP-05-2017-0045

Meier, A. (2021). Studying problems, not problematic usage: Do mobile checking habits increase procrastination and decrease well-being? *Mobile Media & Communication*, 20501579211029326. https://doi.org/10.1177/20501579211029326

Meier, A., & Krause, H.-V. (2023). Does Passive Social Media Use Harm Well-Being?: An Adversarial Review. *Journal of Media Psychology*, *35*(3), 169–180. https://doi.org/10.1027/1864-1105/a000358

Meier, A., & Reinecke, L. (2021). Computer-Mediated Communication, Social Media, and Mental Health: A Conceptual and Empirical Meta-Review. *Communication Research*, *48*(8), 1182–1209. https://doi.org/10.1177/0093650220958224

Molaib, K. M., Sun, X., Ram, N., Reeves, B., & Robinson, T. N. (2025). Agreement between self-reported and objectively measured smartphone use among adolescents and adults. *Computers in Human Behavior Reports*, *17*, 100569. https://doi.org/10.1016/j.chbr.2024.100569

Montanez, G. D., White, R. W., & Huang, X. (2014). Cross-Device Search. *Proceedings of the 23rd ACM International Conference on Conference on Information and Knowledge Management, CIKM '14*, 1669–1678. https://doi.org/10.1145/2661829.2661910

Most, J., Tosti, V., Redman, L. M., & Fontana, L. (2017). Calorie restriction in humans: An update. *Ageing Research Reviews*, *39*, 36–45. https://doi.org/10.1016/j.arr.2016.08.005

Mullan, K., & Wajcman, J. (2019). Have Mobile Devices Changed Working Patterns in the 21st Century? A Time-diary Analysis of Work Extension in the UK. *Work, Employment and Society*, *33*(1), 3–20. https://doi.org/10.1177/0950017017730529

Murphy, S. L., Van Gaeveren, K., De Segovia Vicente, D., & Vanden Abeele, M. (2025). Analyzing the Affective Consequences of Normal Sleep Fluctuations: A Multiverse Investigation Using Experience Sampling Data. *Collabra: Psychology*, *11*(1), 138643. https://doi.org/10.1525/collabra.138643

Nagy, P., Eschrich, J., & Finn, E. (2021). Time hacking: How technologies mediate time. *Information, Communication & Society*, *24*(15), 2229–2243. https://doi.org/10.1080/1369118X.2020.1758743

Naldi, A., Vallelonga, F., Liberto, A. D., Cavallo, R., Agnesone, M., Gonella, M., Sauta, M. D., Lochner, P., Tondo, G., Bragazzi, N. L., Botto, R., & Leombruni, P. (2021). COVID-19 pandemic-related anxiety, distress and burnout: Prevalence and associated factors in healthcare workers of North-West Italy. *BJPsych Open*, *7*(1), e27. https://doi.org/10.1192/bjo.2020.161

Nguyen, M. H. (2021). Managing Social Media Use in an "Always-On" Society: Exploring Digital Wellbeing Strategies That People Use to Disconnect. *Mass Communication and Society*, *24*(6), 795–817. https://doi.org/10.1080/15205436.2021.1979045

Nguyen, M. H. (2023). "Maybe I should get rid of it for a while…": Examining motivations and challenges for social media disconnection. *The Communication Review*, *26*(2), 125–150. https://doi.org/10.1080/10714421.2023.2195795

Nisafani, A. S., Kiely, G., & Mahony, C. (2020). Workers' technostress: A review of its causes, strains, inhibitors, and impacts. *Journal of Decision Systems*, *29*, 243–258. https://doi.org/10.1080/12460125.2020.1796286

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

Olteanu, A., Castillo, C., Diaz, F., & Kıcıman, E. (2019). Social Data: Biases, Methodological Pitfalls, and Ethical Boundaries. *Frontiers in Big Data*, *2*. https://www.frontiersin.org/articles/10.3389/fdata.2019.00013

Orben, A., & Przybylski, A. K. (2019). The association between adolescent well-being and digital technology use. *Nature Human Behaviour*, *3*(2), 173–182. https://doi.org/10.1038/s41562-018-0506-1

Oulasvirta, A., Rattenbury, T., Ma, L., & Raita, E. (2012). Habits make smartphone use more pervasive. *Personal and Ubiquitous Computing*, *16*(1), 105–114. https://doi.org/10.1007/s00779-011-0412-2

Pan, Y.-C., Lin, H.-H., Chiu, Y.-C., Lin, S.-H., & Lin, Y.-H. (2019). Temporal stability of smartphone use data: Determining fundamental time unit and independent cycle. *JMIR mHealth and uHealth*, *7*(3), e12171. https://doi.org/10.2196/12171

Pankowska, P., Cernat, A., Keusch, F., & Bach, R. (2025). Using hidden Markov models to assess and correct for measurement error in digital trace data. *Communication Methods and Measures*, *0*(0), 1–25. https://doi.org/10.1080/19312458.2025.2573265

Parry, D. A., & Coetzee, B. (2025). Do mindsets really matter? A second look at how perceptions of social media experiences relate to well-being. *Journal of Computer-Mediated Communication*, *30*(4), zmaf011. https://doi.org/10.1093/jcmc/zmaf011

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, 11), 1535–1547. https://doi.org/10.1038/s41562-021-01117-5

Parry, D. A., Fisher, J. T., Mieczkowski, H., Sewall, C. J. R., & Davidson, B. I. (2022). Social media and well-being: A methodological perspective. *Current Opinion in Psychology*, *45*, 101285. https://doi.org/10.1016/j.copsyc.2021.11.005

Parry, D. A., & Klingelhoefer, J. (2025, July 22). *Traces as Data*. https://doi.org/10.31234/osf.io/6nctf_v1

Parry, D. A., & Toth, R. (2025). Extracting Meaningful Measures of Smartphone Usage from Android Event Log Data: A Methodological Primer. *Computational Communication Research*, *7*(1), 1. https://doi.org/10.5117/CCR2025.1.8.PARR

Peng, T.-Q., Zhou, Y., & Zhu, J. J. H. (2020). From Filled to Empty Time Intervals: Quantifying Online Behaviors with Digital Traces. *Communication Methods and Measures*, *14*(4), 219–238. https://doi.org/10.1080/19312458.2020.1812556

Peng, Y., & Tullis, J. G. (2021). Dividing attention impairs metacognitive control more than monitoring. *Psychonomic Bulletin & Review*, *28*(6), 2064–2074. https://doi.org/10.3758/s13423-021-01950-9

Peters, H., Bayer, J. B., Matz, S. C., Chi, Y., Vaid, S. S., & Harari, G. M. (2024). Social media use is predictable from app sequences: Using LSTM and transformer neural networks to model habitual behavior. *Computers in Human Behavior*, *161*, 108381. https://doi.org/10.1016/j.chb.2024.108381

Piszczek, M. M. (2017). Boundary control and controlled boundaries: Organizational expectations for technology use at the work–family interface. *Journal of Organizational Behavior*, *38*(4), 592–611. https://doi.org/10.1002/job.2153

Pratap, A., Atkins, D. C., Renn, B. N., Tanana, M. J., Mooney, S. D., Anguera, J. A., & Areán, P. A. (2019). The accuracy of passive phone sensors in predicting daily mood. *Depression and Anxiety*, *36*(1), 72–81. https://doi.org/10.1002/da.22822

Pritchard, K., & Symon, G. (2023). Urgency at work: Trains, time and technology. *New Technology, Work and Employment*, *38*(3), 453–471. https://doi.org/10.1111/ntwe.12269

Prommer, E. (2019). Polychronicity During Simultaneity: Mediated Time and Mobile Media. In M. Hartmann, E. Prommer, K. Deckner, & S. O. Görland (Eds.), *Mediated Time: Perspectives on Time in a Digital Age* (pp. 299–319). Springer International Publishing. https://doi.org/10.1007/978-3-030-24950-2_15

Reeves, B., Ram, N., Robinson, T. N., Cummings, J. J., Giles, C. L., Pan, J., Chiatti, A., Cho, M., Roehrick, K., Yang, X., Gagneja, A., Brinberg, M., Muise, D., Lu, Y., Luo, M., Fitzgerald, A., & Yeykelis, L. (2021). Screenomics: A framework to capture and analyze personal life experiences and the ways that technology shapes them. *Human-Computer Interaction*. https://doi.org/10.1080/07370024.2019.1578652

Reinecke, L., & Hofmann, W. (2016). Slacking off or Winding down? An Experience Sampling Study on the Drivers and Consequences of Media Use for Recovery versus Procrastination. *Human Communication Research*, *42*(3), 441–461. https://doi.org/10.1111/hcre.12082

Reinecke, L., Klimmt, C., Meier, A., Reich, S., Hefner, D., Knop-Huelss, K., Rieger, D., & Vorderer, P. (2018). Permanently online and permanently connected: Development and validation of the Online Vigilance Scale. *PLOS ONE*, *13*(10), e0205384. https://doi.org/10.1371/journal.pone.0205384

Rioja, K., Cekic, S., Bavelier, D., & Baumgartner, S. E. (2023). Unraveling the link between media multitasking and attention across three samples. *Technology, Mind, and Behavior*, *4*(2), 134–150. https://doi.org/10.1037/tmb0000106

Robinson, J. P. (2013). Americans Less Rushed But No Happier: 1965–2010 Trends in Subjective Time and Happiness. *Social Indicators Research*, *113*(3), 1091–1104. https://doi.org/10.1007/s11205-012-0133-6

Rodman, A. M., Burns, J. A., Cotter, G. K., Ohashi, Y.-G. B., Rich, R. K., & McLaughlin, K. A. (2024). Within-Person Fluctuations in Objective Smartphone Use and Emotional Processes During Adolescence: An Intensive Longitudinal Study. *Affective Science*, *5*(4), 332–345. https://doi.org/10.1007/s42761-024-00247-z

Rosa, H. (2013). *Social acceleration: A new theory of modernity* (J. Trejo-Mathys, Trans.). Columbia University Press.

Rosa, H. (2017). Dynamic Stabilization, the Triple A. Approach to the Good Life, and the Resonance Conception. *Questions de Communication*, (31), 437–456. https://doi.org/10.4000/questionsdecommunication.11228

Ross, M. Q., Bayer, J., Rhee, L., Potti, I., & Chang, Y.-J. (2023). Tracking the temporal flows of mobile communication in daily life. *New Media & Society*, *25*(4), 732–755. https://doi.org/10.1177/14614448231158646

Rosseel, Y. (2012). Lavaan: An R package for structural equation modeling. *Journal of Statistical Software*, *48*(2), 1–36. https://doi.org/10.18637/jss.v048.i02

Ryan, R. M., & Deci, E. L. (2008). From Ego Depletion to Vitality: Theory and Findings Concerning the Facilitation of Energy Available to the Self. *Social and Personality Psychology Compass*, *2*(2), 702–717. https://doi.org/10.1111/j.1751-9004.2008.00098.x

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

Saito, R., Kuboyama, T., & Yasuda, H. (2015). User behaviour modelling by abstracting low-level window transition logs. *International Journal of Computational Science and Engineering*. https://www.inderscienceonline.com/doi/10.1504/IJCSE.2015.072648

Santarius, T., & Bergener, J. (2020). Does the Use of ICT speed up the Pace of Life? *Proceedings of the 7th International Conference on ICT for Sustainability, ICT4S2020*, 107–115. https://doi.org/10.1145/3401335.3401642

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

Schaufeli, W. B., Desart, S., & De Witte, H. (2020). Burnout Assessment Tool (BAT)-Development, Validity, and Reliability. *International Journal of Environmental Research and Public Health*. https://doi.org/10.3390/IJERPH17249495

Schenkel, K., Nussbeck, F. W., Kerkhoff, D., Scholz, U., Keller, J., & Radtke, T. (2024). Digital generation: The association between daily smartphone app use and well-being. *Behaviour & Information Technology*, *0*(0), 1–18. https://doi.org/10.1080/0144929X.2024.2406255

Schoch, D., Chan, C., Wagner, C., & Bleier, A. (2023, October 4). *Computational Reproducibility in Computational Social Science*. https://doi.org/10.48550/arXiv.2307.01918

Schoedel, R., Oldemeier, M., Bonauer, L., & Sust, L. (2022). Systematic Categorisation of 3,091 Smartphone Applications From a Large-Scale Smartphone Sensing Dataset. *Journal of Open Psychology Data*, *10*(1). https://doi.org/10.5334/jopd.59

Sevenhant, R., Stragier, J., De Marez, L., & Schuurman, D. (2022). *Imec Digimeter 2021. Digitale trends in Vlaanderen* (p. 180). www.imec.be/digimeter

Sewall, C. J. R., Bear, T. M., Merranko, J., & Rosen, D. (2020). How psychosocial well-being and usage amount predict inaccuracies in retrospective estimates of digital technology use. *Mobile Media & Communication*, *8*(3), 379–399. https://doi.org/10.1177/2050157920902830

Sewall, C. J. R., Goldstein, T. R., Wright, A. G. C., & Rosen, D. (2022). Does Objectively Measured Social-Media or Smartphone Use Predict Depression, Anxiety, or Social Isolation Among Young Adults? *Clinical Psychological Science*, 21677026221078309. https://doi.org/10.1177/21677026221078309

Shaw, H., Ellis, D. A., Kendrick, L.-R., Ziegler, F., & Wiseman, R. (2016). Predicting Smartphone Operating System from Personality and Individual Differences. *Cyberpsychology, Behavior, and Social Networking*, *19*(12), 727–732. https://doi.org/10.1089/cyber.2016.0324

Sheldon, K. M., Ryan, R., & Reis, H. T. (1996). What Makes for a Good Day? Competence and Autonomy in the Day and in the Person. *Personality and Social Psychology Bulletin*, *22*(12), 1270–1279. https://doi.org/10.1177/01461672962212007

Siebers, T., Beyens, I., & Valkenburg, P. M. (2024). The effects of fragmented and sticky smartphone use on distraction and task delay. *Mobile Media & Communication*, *12*(1), 45–70. https://doi.org/10.1177/20501579231193941

Smith, M. R., Chai, R., Nguyen, H. T., Marcora, S. M., & Coutts, A. J. (2019). Comparing the Effects of Three Cognitive Tasks on Indicators of Mental Fatigue. *The Journal of Psychology*, *153*(8), 759–783. https://doi.org/10.1080/00223980.2019.1611530

Steyvers, M., Hawkins, G. E., Karayanidis, F., & Brown, S. D. (2019). A large-scale analysis of task switching practice effects across the lifespan. *Proceedings of the National Academy of Sciences*, *116*(36), 17735–17740. https://doi.org/10.1073/pnas.1906788116

Sussman, R. F., & Sekuler, R. (2022). Feeling rushed? Perceived time pressure impacts executive function and stress. *Acta Psychologica*, *229*, 103702. https://doi.org/10.1016/j.actpsy.2022.103702

team, T. pandas development. (2020). *Pandas-dev/pandas: Pandas* (Version 1.5.3) [Computer software]. Zenodo. https://doi.org/10.5281/zenodo.3509134

The Polars Development Team. (2025). *Polars* (Version 1.35.1) [Computer software]. https://pola.rs/

Thomas, F. (2022). A Methodological Framework for Analyzing the Appearance and Duration of Media Effects. *Journal of Communication*, jqac013. https://doi.org/10.1093/joc/jqac013

Thulin, E., & Vilhelmson, B. (2022). Pacesetters in contemporary telework: How smartphones and mediated presence reshape the time–space rhythms of daily work. *New Technology, Work and Employment*, *37*(2), 250–269. https://doi.org/10.1111/ntwe.12224

Thulin, E., Vilhelmson, B., & Johansson, M. (2019). New Telework, Time Pressure, and Time Use Control in Everyday Life. *Sustainability*, *11*(11, 11), 3067. https://doi.org/10.3390/su11113067

Ting, P.-H., Hwang, J.-R., Doong, J.-L., & Jeng, M.-C. (2008). Driver fatigue and highway driving: A simulator study. *Physiology & Behavior*, *94*(3), 448–453. https://doi.org/10.1016/j.physbeh.2008.02.015

Tkaczyk, M., Tancoš, M., Smahel, D., Elavsky, S., & Plhák, J. (2024). (In)accuracy and convergent validity of daily end-of-day and single-time self-reported estimations of smartphone use among adolescents. *Computers in Human Behavior*, *158*, 108281. https://doi.org/10.1016/j.chb.2024.108281

Toth, R., Parry, D. A., & Klingelhoefer, J. (2025, November 13). *Somebody's (Still) Watching Me: Reactivity to Smartphone Logging and Experience Sampling* (xt24p_v1). https://doi.org/10.31235/osf.io/xt24p_v1 (Pre-published)

Toth, R., & Trifonova, T. (2021). Somebody's Watching Me: Smartphone Use Tracking and Reactivity. *Computers in Human Behavior Reports*, *4*, 100142. https://doi.org/10.1016/j.chbr.2021.100142

Tseng, F.-C., Chiou, Z.-Y., Chuang, H.-H., Su, L.-T., Lin, Y.-H., Lin, Y.-R., Lee, Y.-C., Wang, P.-J., Chen, U.-D., & Chang, Y.-J. (2023). Multiple Device Users' Actual and Ideal Cross-Device Usage for Multi-Stage Notification-Interactions: An ESM Study Addressing the Usage Gap and Impacts of Device Context. *Proceedings of the 2023 CHI Conference on Human Factors in Computing Systems*, 1–15. https://doi.org/10.1145/3544548.3580731

Twenge, J. M., & Farley, E. (2021). Not all screen time is created equal: Associations with mental health vary by activity and gender. *Social Psychiatry and Psychiatric Epidemiology*, *56*(2), 207–217. https://doi.org/10.1007/s00127-020-01906-9

Twenge, J. M., Joiner, T. E., Rogers, M. L., & Martin, G. N. (2018). Increases in Depressive Symptoms, Suicide-Related Outcomes, and Suicide Rates Among U.S. Adolescents After 2010 and Links to Increased New Media Screen Time. *Clinical Psychological Science*, *6*(1), 3–17. https://doi.org/10.1177/2167702617723376

Vaid, S. S., Kroencke, L., Roshanaei, M., Talaifar, S., Hancock, J. T., Back, M. D., Gosling, S. D., Ram, N., & Harari, G. M. (2024). Variation in social media sensitivity across people and contexts. *Scientific Reports*, *14*(1), 6571. https://doi.org/10.1038/s41598-024-55064-y

Valkenburg, P. M. (2022). Social media use and well-being: What we know and what we need to know. *Current Opinion in Psychology*, *45*, 101294. https://doi.org/10.1016/j.copsyc.2021.12.006

Valkenburg, P. M., Beyens, I., de Vaate, N. B., Janssen, L., & van der Wal, A. (2024). Person-Specific Media Effects. In T. Araujo & P. Neijens (Eds.), *Communication Research into the Digital Society* (pp. 233–246). Amsterdam University Press. https://doi.org/10.2307/jj.11895525.17

Valkenburg, P. M., Meier, A., & Beyens, I. (2022). Social media use and its impact on adolescent mental health: An umbrella review of the evidence. *Current Opinion in Psychology*, *44*, 58–68. https://doi.org/10.1016/j.copsyc.2021.08.017

Valkenburg, P. M., Pouwels, J. L., Beyens, I., Driel, I. I. van, & Keijsers, L. (2021). Adolescents' Social Media Experiences and Their Self-Esteem: A Person-Specific Susceptibility Perspective. *Technology, Mind, and Behavior*, *2*(2). https://doi.org/10.1037/tmb0000037

Vallat, R. (2018). Pingouin: Statistics in python. *Journal of Open Source Software*, *3*(31), 1026. https://doi.org/10.21105/joss.01026

van Atteveldt, W., & Peng, T.-Q. (2018). When Communication Meets Computation: Opportunities, Challenges, and Pitfalls in Computational Communication Science. *Communication Methods and Measures*, *12*(2–3), 81–92. https://doi.org/10.1080/19312458.2018.1458084

Van Bruyssel, S., De Wolf, R., & Vanden Abeele, M. (2026). From bliss to burden: An ethnographic inquiry into how social, material and individual obstacles to digital well-being shape everyday life. *New Media & Society*, *28*(1), 212–229. https://doi.org/10.1177/14614448241288159

Van Gaeveren, K., Murphy, S., de Segovia Vicente, D., & Vanden Abeele, M. (2025). Always On, Always Rushed For Time? Exploring Momentary Associations Between Passively Sensed Smartphone Use, Feeling Rushed, and Perceived Task Juggling. *Mobile Media & Communication*, 20501579251377010. https://doi.org/10.1177/20501579251377010

Vanden Abeele, M. M. P. (2021). Digital Wellbeing as a Dynamic Construct. *Communication Theory*, *31*(4), 932–955. https://doi.org/10.1093/ct/qtaa024

Vanden Abeele, M. M. P., De Wolf, R., & Ling, R. (2018). Mobile Media and Social Space: How Anytime, Anyplace Connectivity Structures Everyday Life. *Media and Communication*, *6*(2), 5–14. https://doi.org/10.17645/mac.v6i2.1399

Vanden Abeele, M. M. P., Murphy, S. L., Lemahieu, L., & Koster, E. H. W. (2025). Methodological considerations for social media intervention studies. *Nature Reviews Psychology*, *4*(9), 603–614. https://doi.org/10.1038/s44159-025-00474-6

Vanden Abeele, M. M. P., & Nguyen, M. H. (2024). Digital media as ambiguous goods: Examining the digital well-being experiences and disconnection practices of Belgian adults. *European Journal of Communication*, *39*(2), 122–144. https://doi.org/10.1177/02673231231201487

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

Vandenbosch, L., Beullens, K., Vanherle, R., & Schreurs, L. (2025). Digital media uses and effects: The contributing roles of time. *Journal of Children and Media*, *19*(1), 71–76. https://doi.org/10.1080/17482798.2024.2438690

Vanluydt, E., van den Eijnden, R., Vonk, L., Putrik, P., van Amelsvoort, T., Delespaul, P., Levels, M., & Huijts, T. (2026). Disconnect To Reconnect: How Variations between Types of Smartphone Bans Influence Students' Well-being and Social Connectedness in Dutch Secondary Education. *Journal of Youth and Adolescence*. https://doi.org/10.1007/s10964-025-02313-6

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

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

Viechtbauer, W., & Constantin, M. (2023). *Esmpack: Functions that facilitate preparation and management of ESM/EMA data* [Manual].

Vorderer, P., Hefner, D., Reinecke, L., & Klimmt, C. (2017). Permanently Online and Permanently Connected: A New Paradigm in Communication Research? In *Permanently Online, Permanently Connected*. Routledge.

Vorderer, P., Hefner, D., Reinecke, L., & Klimmt, C. (Eds.). (2018). *Permanently online, permanently connected: Living and communicating in a POPC world*. Routledge, Taylor & Francis Group.

Vostal, F. (2021). Aceleración social: cinco comentarios "deflacionarios." *Res Publica. Revista de Historia de las Ideas Políticas*, *24*(3, 3), 446–453. https://doi.org/10.5209/rpub.79243

Wajcman, J. (2008). Life in the fast lane? Towards a sociology of technology and time. *The British Journal of Sociology*, *59*(1), 59–77. https://doi.org/10.1111/j.1468-4446.2007.00182.x

Wajcman, J. (2014). *Pressed for Time: The Acceleration of Life in Digital Capitalism*. University of Chicago Press. https://doi.org/10.7208/chicago/9780226196503.001.0001

Wajcman, J. (2015). *Pressed for time: The acceleration of life in digital capitalism*. The University of Chicago Press.

Wang, L., & Miller, L. C. (2020). Just-in-the-Moment Adaptive Interventions (JITAI): A Meta-Analytical Review. *Health Communication*, *35*(12), 1531–1544. https://doi.org/10.1080/10410236.2019.1652388

Warm, J. S., Parasuraman, R., & Matthews, G. (2008). Vigilance Requires Hard Mental Work and Is Stressful. *Human Factors: The Journal of the Human Factors and Ergonomics Society*, *50*(3), 433–441. https://doi.org/10.1518/001872008X312152

Wearden, J. (2016). *The Psychology of Time Perception*. Palgrave Macmillan UK. https://doi.org/10.1057/978-1-137-40883-9

Weiss, H. A., & Bonell, C. (2025). Smartphone use and mental health: Going beyond school restriction policies. *The Lancet Regional Health – Europe*, *51*. https://doi.org/10.1016/j.lanepe.2025.101237

Wickham, H., François, R., Henry, L., Müller, K., & Vaughan, D. (2025). *Dplyr: A grammar of data manipulation* [Computer software]. https://dplyr.tidyverse.org

Williams, J. (2018). *Stand out of our Light: Freedom and Resistance in the Attention Economy* (1st ed.). Cambridge University Press. https://doi.org/10.1017/9781108453004

Williams, K. J., Suls, J., Alliger, G. M., Learner, S. M., & Wan, C. K. (1991). Multiple role juggling and daily mood states in working mothers: An experience sampling study. *Journal of Applied Psychology*, *76*(5), 664–674. https://doi.org/10.1037/0021-9010.76.5.664

Winbush, A., McDuff, D., Hernandez, J., Barakat, A., Jiang, A., Heneghan, C., Nelson, B. W., & Allen, N. B. (2025). Smartphone use in a large US adult population: Temporal associations between objective measures of usage and mental well-being. *Proceedings of the National Academy of Sciences*, *122*(43), e2427311122. https://doi.org/10.1073/pnas.2427311122

Wiradhany, W., Baumgartner, S., & De Bruin, A. (2021). Exploitation–Exploration Model of Media Multitasking. *Journal of Media Psychology*, *33*(4), 169–180. https://doi.org/10.1027/1864-1105/a000303

Wittmann, M. (2009). The inner experience of time. *Philosophical Transactions of the Royal Society B: Biological Sciences*, *364*(1525), 1955–1967. https://doi.org/10.1098/rstb.2009.0003

Wolfers, L. N. (2024). A social constructivist viewpoint of media effects: Extending the social influence model of technology use to media effects. *Communication Theory*, qtae015. https://doi.org/10.1093/ct/qtae015

Yanagida, T. (2023). *Misty: Miscellaneous Functions 'T. Yanagida'* (Version 0.5.0) [Computer software]. https://cran.r-project.org/web/packages/misty/index.html

Ytre-Arne, B., Syvertsen, T., Moe, H., & Karlsen, F. (2020). Temporal ambivalences in smartphone use: Conflicting flows, conflicting responsibilities. *New Media & Society*, *22*(9), 1715–1732. https://doi.org/10.1177/1461444820913561

Summary

English Summary

Smartphones and other digital devices are now part of everyday life. While they are useful, there is concern that being "always on" may come at a cost to our mental well-being. Yet the scientific evidence remains surprisingly inconclusive, in part because most studies still rely on people's own, often inaccurate, estimates of how much time they spend on their devices. This dissertation addresses this gap by combining passive sensing (automatically logging device activity) with experience sampling (asking participants to report their feelings multiple times per day). Together, these methods allow us to study, in real time, what people actually do on their devices and how it makes them feel. The dissertation comprises three empirical studies and one methodological contribution, preceded by a shared introduction and methodology chapter, and concluded by a general discussion.

The methodology chapter details the shared research infrastructure underlying this dissertation. All chapters draw from a large-scale citizen science project conducted at Ghent University in the fall of 2022, in which a total of 1,315 adults were recruited through a citizen-science collaboration with De Standaard newspaper. Over two weeks, participants received six daily questionnaires capturing momentary experiences of well-being, mental fatigue, time pressure, and affect, among others. Simultaneously, the smartphones of 834 Android users were passively logged, with another 106 of them also having their computer use tracked. This combination of over 67,000 questionnaire responses and more than 7 million smartphone events offers an unusually fine-grained view of how digital behavior and well-being interact in daily life. The methodology chapter provides details on the data collection infrastructure, the tools used for passive sensing and experience sampling, and the data processing pipeline that transforms raw device logs into research-ready features. This is an open, reusable reference for researchers conducting similar studies.

The first study (Chapter 3) examines online vigilance – the tendency to be constantly attentive to one's online world. The findings show that moments of heightened online vigilance coincided with greater mental fatigue, and that this effect lingered throughout the day. Interestingly, passively sensed smartphone behavior (e.g., phone checking, notification response speed) were only weakly related to self-reported online vigilance, suggesting that the subjective experience of being "always on" matters more for well-being than what is visible in the behavioral data alone.

The second study (Chapter 4) explores whether mobile communication makes people feel rushed. Results show that email and work communication app use, in particular, were associated with feeling more pressed for time. This was partly explained by the experience of juggling multiple tasks and roles: mobile communication intensified competing demands, accelerating people's sense of time. Social media use was a notable exception, showing no link to feeling

rushed. Some of these associations varied by age, gender, and parenthood, showing that media effects are not one-size-fits-all.

The third study (Chapter 5) challenges the field's near-exclusive focus on smartphones when measuring screen time. Using data from 106 participants who had both their smartphone and computer use logged, the study shows that ignoring computer use underestimates total screen time by an average of 136%. Participants switched between devices about 66 times per day. These cross-device dynamics are invisible to single-device studies. Including these computer-based features altered conclusions about the relationship between digital media use and mood, arguing that the field needs multi-device data collection strategies.

Beyond the three empirical studies, this dissertation includes a methodological contribution (Chapter 6) which addresses a long-standing technical barrier: the inability to passively track screen time on Apple devices. It introduces a new data donation procedure that leverages Apple's built-in Screen Time feature. By enabling synchronization across devices linked to the same Apple ID, participants can donate system-level files from their Mac containing detailed usage across all connected Apple devices. An open-source tool converts these files into researcher-friendly datasets. While the method has limitations (it requires a Mac and captures only four weeks of data), it marks an important step in bridging the gap between Android and Apple device tracking in research.

This dissertation demonstrates that understanding the link between digital media use and well-being requires moving beyond simple screen time metrics. How people experience their connectivity – whether they feel drained, rushed, or happy – matters. The findings consistently show that associations between digital behavior and well-being are small, dynamic, and person-specific. The dissertation contributes new insights into how digital connectivity affects well-being, and advances how researchers can collect, analyze, and document digital traces across devices and platforms.

Nederlandstalige Samenvatting

Smartphones en andere digitale apparaten maken intussen onlosmakelijk deel uit van het dagelijks leven. Hoewel ze nuttig zijn, bestaat de bezorgdheid dat het "altijd online" zijn ten koste kan gaan van ons mentaal welzijn. Toch blijft het wetenschappelijk bewijs verrassend onbeslist, mede doordat de meeste studies nog steeds steunen op de eigen, vaak onnauwkeurige, schattingen van mensen over hun schermtijd. Dit proefschrift pakt deze lacune aan door passieve logging (het automatisch loggen van apparaat activiteit) te combineren met de experience sampling methode (het meermaals per dag bevragen van deelnemers over hun gevoelens). Samen maken deze methoden het mogelijk om in real time te bestuderen wat mensen effectief doen op hun toestellen en hoe zij zich daarbij voelen. Het proefschrift omvat drie empirische studies en één methodologische bijdrage, voorafgegaan door een gedeelde inleiding en een methodologisch hoofdstuk, en afgesloten met een algemene discussie.

Het methodologisch hoofdstuk beschrijft de gedeelde onderzoeksinfrastructuur die aan de basis van dit proefschrift ligt. Alle hoofdstukken komen voort uit een grootschalig burgerwetenschapsproject dat in het najaar van 2022 werd uitgevoerd aan de Universiteit Gent, waarbij in totaal 1.315 volwassenen werden gerekruteerd via een samenwerking met De Standaard. Gedurende twee weken ontvingen de deelnemers zes dagelijkse vragenlijsten die ervaringen van welzijn, mentale vermoeidheid, tijdsdruk en gemoedstoestand in kaart brachten. Tegelijkertijd werden de smartphones van 834 Android-gebruikers passief gelogd, en werd bij 106 van hen ook het computergebruik geregistreerd. Deze combinatie van meer dan 67.000 vragenlijstresponsen en ruim 7 miljoen smartphone-events biedt een gedetailleerd beeld van hoe digitaal gedrag en welzijn in het dagelijks leven samenhangen. Het methodologisch hoofdstuk beschrijft de infrastructuur voor dataverzameling, de gebruikte tools voor passieve logging en experience sampling, en de dataverwerkingspipeline die ruwe logs omzet in variabelen. Het vormt een open en herbruikbare referentie voor onderzoekers die vergelijkbare studies uitvoeren.

De eerste studie (Hoofdstuk 3) onderzoekt online vigilance – de neiging om voortdurend aandachtig te zijn voor de eigen online wereld. De bevindingen tonen aan dat momenten van verhoogde online waakzaamheid samenvielen met grotere mentale vermoeidheid, en dat dit effect doorwerkte gedurende de rest van de dag. Opvallend is dat passief geregistreerd smartphonegedrag (zoals het checken van de telefoon en de reactiesnelheid op notificaties) slechts zwak samenhing met zelf gerapporteerde online waakzaamheid. Dit suggereert dat de subjectieve ervaring van het "altijd online" zijn meer uitmaakt voor welzijn dan wat zichtbaar is in de gedragsdata.

De tweede studie (Hoofdstuk 4) gaat na of mobiele communicatie mensen het gevoel geeft gehaast te zijn. De resultaten tonen dat vooral het gebruik van e-mail- en werkcommunicatie-apps samenhing met een sterker gevoel van tijdsdruk. Dit werd deels verklaard door de ervaring meerdere taken en rollen tegelijk te moeten jongleren: mobiele communicatie versterkte concurrerende eisen, waardoor het tijdsgevoel van mensen versnelde. Het gebruik van sociale media vormde een opvallende uitzondering en vertoonde geen verband met het gevoel gehaast te zijn. Sommige van deze verbanden verschilden naargelang leeftijd, geslacht en ouderschap, wat aantoont dat media-effecten niet voor iedereen gelijk zijn.

De derde studie (Hoofdstuk 5) stelt de bijna exclusieve focus op smartphones bij het meten van schermtijd in vraag. Op basis van gegevens van 106 deelnemers bij wie zowel het smartphone- als het computergebruik werd gelogd, toont de studie aan dat het negeren van computergebruik de totale schermtijd gemiddeld met 136% onderschat. Deelnemers wisselden dagelijks ongeveer 66 keer tussen apparaten. Deze 'cross-device' dynamieken zijn onzichtbaar in studies die slechts één toestel registreren. Het opnemen van computer gerelateerde variabelen veranderde de conclusies over het verband tussen digitaal mediagebruik en stemming, en pleit ervoor dat het veld nood heeft aan dataverzamelingsstrategieën die meerdere apparaten registreren.

Naast de drie empirische studies bevat dit proefschrift een methodologische bijdrage (Hoofdstuk 6) die een langdurige technische barrière aanpakt: de onmogelijkheid om schermtijd op Apple-apparaten passief te registreren. Het introduceert een nieuwe datadonatieprocedure die gebruikmaakt van Apple's ingebouwde Schermtijd-functie. Door synchronisatie in te schakelen tussen apparaten die aan dezelfde Apple ID zijn gekoppeld, kunnen deelnemers systeembestanden van hun Mac doneren die gedetailleerde gebruiksgegevens bevatten van alle verbonden Apple-apparaten. Een open-sourcetool zet deze bestanden om naar bruikbare datasets. Hoewel de methode beperkingen kent (ze vereist een Mac en registreert slechts vier weken aan gegevens), vormt ze een belangrijke stap in het overbruggen van de kloof tussen Android- en Apple-apparaat registratie in onderzoek.

Dit proefschrift toont aan dat het begrijpen van het verband tussen digitaal mediagebruik en welzijn vereist dat we voorbij eenvoudige schermtijdmaten kijken. Hoe mensen hun connectiviteit ervaren – of ze zich uitgeput, gehaast of gelukkig voelen – doet ertoe. De bevindingen tonen consistent aan dat verbanden tussen digitaal gedrag en welzijn klein, dynamisch en persoon specifiek zijn. Het proefschrift levert nieuwe inzichten in hoe digitale connectiviteit welzijn beïnvloedt, en draagt bij aan de manier waarop onderzoekers digitale sporen (of 'traces') kunnen verzamelen, analyseren en documenteren over apparaten en platformen heen.