5-11-2023

DANGEROUS HABIT OR USEFUL ROUTINE? DEVELOPING THEORY-BASED MEASURES FOR THE INTENDED AND UNINTENDED CONSEQUENCES OF STATE-TRACKING HABITS

Arvid Alexander Eichner
Goethe University Frankfurt, eichner@its.uni-frankfurt.de

Nicholas Valentin Lingnau
Goethe University Frankfurt, lingnau@its.uni-frankfurt.de

Patrick Felka
Goethe University Frankfurt, felka@wiwi.uni-frankfurt.de

Vanessa Kohn
Goethe University Frankfurt am Main, kohn@its.uni-frankfurt.de

Roland Holten
Goethe University Frankfurt, holten@its.uni-frankfurt.de

See next page for additional authors

Authors

Arvid Alexander Eichner, Nicholas Valentin Lingnau, Patrick Felka, Vanessa Kohn, Roland Holten, and Oliver Hinz

DANGEROUS HABIT OR USEFUL ROUTINE? DEVELOPING THEORY-BASED MEASURES FOR THE INTENDED AND UNINTENDED CONSEQUENCES OF STATE-TRACKING HABITS

Research Paper

Arvid Alexander Eichner, Goethe University Frankfurt, Germany, eichner@its.uni-frankfurt.de

Nicholas Valentin Lingnau, Goethe University Frankfurt, Germany, lingnau@its.uni-frankfurt.de

Patrick Felka, Goethe University Frankfurt, Germany, felka@wiwi.uni-frankfurt.de

Vanessa Kohn, Goethe University Frankfurt, Germany, kohn@its.uni-frankfurt.de

Roland Holten, Goethe University Frankfurt, Germany, holten@its.uni-frankfurt.de

Oliver Hinz, Goethe University Frankfurt, Germany, hinz@wiwi.uni-frankfurt.de

Abstract

According to the theory of IT-mediated state-tracking, the intended and unintended consequences of constantly checking digital devices can be judged by the resulting problem of attention and by whether the checking led to information that served the individual's enduring goal. While this perspective offers numerous benefits over the common practice of labeling excessive information technology use as addiction, as of yet, the constructs of problem of attention (PoA) and service to enduring goal (SEG) lack empirical measures. Thus, this paper develops measurement scales for the constructs PoA and SEG following an established construct development methodology. We evaluate the measures' validity and reliability and demonstrate that PoA and SEG differ from existing similar concepts. With the help of our newly developed constructs the quality of constant checking habits can be assessed which enables future studies to scrutinize the theorized preventive role of self-control in the context of smartphone habits.

Keywords: Constant Checking, IT-mediated State-Tracking, Habits, Construct Development.

1 Introduction

Even in situations where their attention is required, people regularly check their smartphones for new information, missed messages, or updates on social media platforms. This behavior, called "constant checking", is a form of excessive smartphone use (Gerlach and Cenfetelli, 2020). When smartphone use becomes excessive or when users fail to control the urge to check for new information constantly, it can lead to stress (Condliffe, 2017), work-home conflicts (Boswell and Olson-Buchanan, 2007), addiction (Salehan and Negahban, 2013; Wu et al., 2013), or depression (Lee et al., 2014). Recently, many researchers have attended to this matter by dedicating their studies to uncovering antecedents and consequences of excessive smartphone use. As a result, many different constructs, such as problematic use (Cheng and Meng, 2021; Elhai et al., 2018; Regan et al., 2020; Zerach, 2020), excessive use (Aranda and Baig, 2018; Holte and Ferraro, 2020; Loid et al., 2020), compulsive smartphone use (Chen et al., 2017; Wang et al., 2014; Wang and Lee, 2020; Zhang et al., 2014), and habitual smartphone use (Van Deursen et al., 2015; Roffarello and De Russis, 2021) have emerged. The predominant label to describe excessive use of IT however, has been that of IT addiction, or, specific to the context of smartphone use, smartphone addiction (Van Deursen et al., 2015; Lan et al., 2018; Lin et al., 2017).

The common application of the IT addiction perspective to study excessive smartphone use does, however, bear some problems. First, many studies use the terms problematic, addictive, and compulsive interchangeably to describe general problematic use, failing to discriminate between established behavioral phenomena (Wang et al., 2014). Second, the labels used to address or study excessive IT use (e.g., labeling someone an addict) also translate to unique countermeasures and potentially far-reaching consequences for the individuals involved (Gerlach and Cenfetelli, 2020). Third, many researchers have raised construct validity concerns regarding IT addiction (Morahan-Martin, 2005; Shaffer et al., 2000) and it is often applied in the form of an imprecise umbrella term labeling IT use that is high in extent as "IT addiction" (Gerlach and Cenfetelli, 2020; Wang and Lee, 2020). Fourth, choosing the smartphone as the unit of analysis instead of accounting for the different features or applications used is imprecise at best (Wang and Lee, 2020). Single-IS theories might be incomplete, could mispresent the phenomenon, or might be non-existent or not applicable for phenomena that entail the use of multiple IS (Gerlach and Cenfetelli, 2021). Although the extent of smartphone use can easily and precisely be measured (time spent actively using the device, known as screen time), individuals use many different applications with a wide variety of purposes ranging from entertainment and games to social networking and utilities. Some of these categories, such as social networking, chatting, or gaming have been shown to be more strongly linked to compulsive behavior (Meerkerk et al., 2010; van Rooij et al., 2010). Scholars have therefore suggested that, rather than general smartphone use, the use of specific applications should be examined (Wang and Lee, 2020).

Last, without accounting for individual factors, e.g., the often situation-dependent goal of smartphone or application use, it is impossible to make judgments about said use. For example, it would be wrong to hold an influencer, who earns money by spending several hours every day on their social media app of choice to the same standard as someone who is trying to study for an exam but who distracts themselves by using a certain app. Rather than focusing solely on the extent of smartphone use, e.g., by studying its absence ("digital detox") or measuring the impact of different antecedents to excessive smartphone use, the field should focus on uncovering and quantifying the actual problem in problematic smartphone use, refraining from normative judgment and imprecise umbrella terms which are applied regardless of contextual and individual factors (Gerlach and Cenfetelli, 2020, 2021; Panova and Carbonell, 2018; Wang and Lee, 2020).

In summary, the field is plagued by immense theoretical ambiguity and imprecise measurement of constructs. Too many studies follow the general normative assumption of "too much smartphone use is bad for you", a picture famously painted by the popular press (Gerlach and Cenfetelli, 2020). The recently proposed model of IT-mediated state-tracking promises to resolve the issues mentioned above. It offers a less extreme explanation of "constant checking" behavior, a form of excessive smartphone use, by suggesting that this primarily purposeful behavior can be explained by information-seeking habits motivated by individuals' enduring goals, rather than by addiction (Gerlach and Cenfetelli, 2020). It is important to note that, while providing an answer, the theory of IT-mediated state-tracking cannot replace the concept of IT addiction altogether. Rather it offers an alternative explanation to one form of excessive IT use, which is commonly mislabeled as IT addiction (Gerlach and Cenfetelli, 2020). IT users can indeed get addicted to certain behaviors conducted by means of an IT artifact, with their smartphone working as a mediator (Serenko and Turel, 2020). Therefore, despite the theoretical ambiguity outlined above, we should be cautious not to dismiss prior contributions in these fields. The theory of IT-mediated state-tracking accounts for both intended and unintended consequences of "constant checking" habits in the form of the habits' service to enduring goals (SEG) and resulting problems of attention (PoA). These consequences play a crucial role, allowing researchers and individuals to distinguish between state-tracking habits with predominantly intended or unintended consequences. To date, the proposed constructs lack an empirical measure. Therefore, this study aims to develop valid and reliable constructs for both PoA and SEG using the construct development methodology proposed by Lewis et al. (2005) consisting of the three steps (1) domain definition, (2) instrument construction, and (3) evaluation of measurement properties. In the following section, we describe our application of the method and outline its results. We theoretically and empirically show that our newly developed constructs discriminate from similar existing constructs. Last, we discuss our contributions to theory and practice in Section 7 before concluding by highlighting this study's limitations and implications for future research in Section 8.

2 Theoretical Background

The theory of IT-mediated state-tracking states that individuals use IT habitually "to seek information that closes the gap between their knowledge about a real-world domain's state and its actual state" (Gerlach and Cenfetelli, 2020, p. 1706). Initially motivated by so-called enduring goals (e.g., maintaining social relationships, doing well at one's job, or being perceived as competent), individuals develop a need to stay up to date, which, in turn, leads to the development of state-tracking habits. In the behavioral literature, the term habit is used inconsistently with studies portraying habits as either (1) a form of behavior, (2) a tendency to engage in behavior, or (3) the automaticity of responses. The first two views lead to logical inconsistencies as (1) "a habit cannot be both the behavior and the cause of a behavior" (Maddux, 1997, p. 336) and (2) stating that habits are performed due to a tendency is truistic. (Gardner, 2015). Therefore, consistent with the definition used by Gerlach and Cenfetelli (2020), we conceptualize habits to describe a form of automaticity or more specifically "a process by which a stimulus automatically generates an impulse towards action, based on learned stimulus-response associations" (Gardner, 2015, p. 280).

In order to account for the ambivalence of IS use (Qahri-Saremi and Turel, 2020) and while acknowledging that, without subjective judgment, state-tracking habits cannot be labeled "good" or "bad", the authors instead distinguish between intended and unintended consequences of state-tracking habits (Gerlach and Cenfetelli, 2020). As a result, each instance of use for every single app can be evaluated separately based on the outcome it creates. The intended consequences of state-tracking habits are represented by the service to one's enduring goal. In other words, individual state-tracking habits can be judged based on whether the "situational value of information" derived during the execution of said habit serves the enduring goal of its users (Gerlach and Cenfetelli, 2020, p. 1718). For example, someone might have the enduring goal of being up-to-date with current events. As a result, they developed the need to stay up to date with articles in a news app on their smartphone. After repeatedly checking the news app in stable contexts, e.g., while on their way to work, a habit is formed (De Guinea and Markus, 2009; Lally and Gardner, 2013; Limayem et al., 2007; Polites and Karahanna, 2013). If, during the execution of the habit, the user reads an article that provides information on current events, e.g., the outcome of a recent election, the habit provided a service to that user's enduring goal. It would, however, be of no service if the article was not providing any information on current events, e.g., presenting summer gardening tips for the balcony. At first glance, SEG seems similar, if not identical to the well-established concept of perceived usefulness (PU) proposed by Davis (1989) as part of the technology acceptance model. While the two concepts are indeed similar, we are able to show that these concepts are distinct and highlight the key differences in our related variable analysis in Section 6.3.

Regardless of their situational value, constant checking habits can also lead to unintended consequences when they are performed in situations where they prevent individuals from sustaining focus on stimuli that require such focus (Gerlach and Cenfetelli, 2020). In the theory of IT-mediated state-tracking, these unintended consequences of state-tracking habits are summarized under the problem of attention. For example, when attending a work meeting where one's attention is required, checking emails on the smartphone might prevent one from focusing on the meeting at the same time. In this situation the checking of one's mail app would take away from the attention focused on the work meeting, resulting in a problem of attention. Similar to SEG and PU, PoA shares similarities with established concepts such as compulsive smartphone use (CSU) (Zhang et al., 2014) and habitual smartphone use (HU) (Van Deursen et al., 2015). However, we show that PoA discriminates from these related concepts both theoretically and empirically in the related variable analysis, see Section 6.3.

Through PoA and SEG, state-tracking habits or, more loosely, the observable phenomenon of "constant checking" can be evaluated on an app-level based on the resulting problem of attention and the service to an individual's enduring goal. The need for this granular assessment becomes apparent when considering that individuals exert self-control to resist the urge to perform a state-tracking habit once it is triggered by an external cue (Gerlach and Cenfetelli, 2020). While there is no empirical evidence for this assumption yet, acting rationally, individuals would only exert self-control for state-tracking habits that either provide no value to their enduring goal, cause a problem of attention, or a combination of both. Empirically validating this and related assumptions regarding the role of self-control in preventing the unintended manifestations of smartphone use necessitates a reliable way to measure PoA and SEG. Hence, in the following section, we develop valid and reliable constructs for both PoA and SEG using the construct development methodology proposed by Lewis et al. (2005).

3 Construct Development Methodology

Construct development aims to create valid, precise, and relevant measures of underlying latent constructs (Clark and Watson, 2019; Lewis et al., 2005). Ultimately, all psychological constructs can be operationalized as scales. This endeavor is justified when a new construct is considered to be sufficiently important and distinct from existing constructs (Clark and Watson, 2019). As highlighted in the previous chapter, both problem of attention and service to enduring goal currently lack a suitable data collection instrument. This restrains the advancement of knowledge on the intended and unintended consequences of state-tracking habits. To be able to accurately assess these phenomena in practice, we need to develop a suitable measure. An influential methodological framework for construct development in the field of marketing was postulated by Churchill (1979). Building on this basic framework, (Lewis et al., 2005) have developed the first rigorous and detailed guide on construct development specifically for the Information Systems (IS) field. Numerous IS studies have successfully used this methodological guide to develop constructs and validate their measurement instruments. It is considered good practice to adapt the proposed methodology to a specific research context as prioritization increases efficiency (Lewis et al., 2005). Following this guide, the creation of a new measurement instrument can be divided into three stages, each having distinct deliverables.

The first stage establishes the domain of the idea in three steps. First, the premise (i.e., the purpose) of a construct is specified. Second, the construct is conceptually defined and third, the construct dimensions and respective item statements are specified. In the second stage, a pre-test initiates the process of successive improvements to the instrument. Empirical feedback on the questionnaire is collected from participants who are familiar with the construct. The instrument is improved according to the received suggestions, before being administered to participants similar to the target population for a pilot test. Once more the comments and suggestions by the participants are used to adjust the instrument. The third stage encompasses both exploratory and confirmatory assessments of collected data to validate the measurement instrument.

4 Stage I: Domain

The following section describes the first stage of the construct development procedure applied for PoA and SEG.

4.1 Premise

The premise of a construct specifies its purpose or importance. It can be derived from multiple sources, including literature reviews, questionnaires, or interviews (Lewis et al., 2005). For reasons presented in the introduction of this paper, we draw on the work of Gerlach and Cenfetelli (2020), who uncovered the intended and unintended consequences of state-tracking habits, represented by the respective concepts of PoA and SEG through grounded theory, and formally define the premises.

First, in order to judge the unintended consequences of state-tracking habits, we need to measure the resulting problem of attention. Especially when studying the relationship between self-control and the execution of state-tracking habits proposed in the theoretical model of IT-mediated state-tracking, being able to understand which state-tracking habits create problems of attention and which do not is crucial.

Once individuals are sufficiently motivated to change their behavior they start to exert self-control (Quinn et al., 2010; Wood and Neal, 2007, 2009) or use self-control strategies (Adriaanse et al., 2014; Gillebaart et al., 2016; Gillebaart and de Ridder, 2015; Inzlicht et al., 2021). They might form the behavioral intention to change their behavior based on their attitude and subjective norms, as the theory of reasoned actions proposes (Fishbein and Ajzen, 1980). Conversely, individuals likely won't see the need to exert self-control in order to prevent the execution of a state-tracking habit that doesn't have any unintended consequences. But when the threshold of sufficient motivation is reached, or whether individuals are introspective enough to reflect on both the intended and unintended consequences, is unclear. In order to answer these questions, we need to be able to quantify the magnitude of these consequences. For example, future studies could measure the intention to exert self-control as a function of the magnitude of the intended and unintended consequences of state-tracking habits. This does, however, require valid and reliable measurements, which will be developed in this paper. Since both the magnitude and manifestation of these problems of attention vary between different state-tracking habits, individuals, and contexts (Gerlach and Cenfetelli, 2020), measuring the magnitude of PoA on an app-level helps evaluate the quality of individual state-tracking habits and prevents imprecise generalizations.

Second, as crucial as it is to account for the magnitude of the unintended consequences of state-tracking habits, it is essential to account for the intended consequences. The intended consequences of a state-tracking habit can be evaluated based on whether the situational value of information derived from a state-tracking habit was of service to that individual's enduring goal or not (Gerlach and Cenfetelli, 2020). Even when habits don't provide any value for their enduring goal, individuals might still act out of habit, since habit is a stronger predictor of future actions than intentions are (Polites and Karahanna, 2013; Wood and Neal, 2009; Wood and Rünger, 2016). They might, however, be inclined to exert self-control to prevent the execution of the habit. In order to further our understanding of these theorized effects, being able to quantify the service to the individual's enduring goal is indispensable.

Taken together, the intended and unintended consequences determine the quality of a state-tracking habit. Treating all state-tracking habits the same, regardless of their value, would lead to wrong conclusions, especially when assessing the role of self-control in preventing their execution.

4.2 Conceptual definition

The conceptual definitions should flow from the identified premises (Lewis et al. 2005). We adapt the definitions of the concepts provided by Gerlach and Cenfetelli (2020).

The constructs are defined as follows: Problems of attention are defined as *situations that* "*occur when an individual does not focus or sustain focus on a stimulus that requires such focus*" (Gerlach and Cenfetelli, 2020, p. 1718).

A state-tracking habits' service to an enduring goal is defined as *the degree to which the* "*information obtained about the current state of a real-world domain would serve th*[*is*] *enduring goal*[]" (Gerlach and Cenfetelli, 2020, p. 1717).

4.3 Dimensions and items

Since the goal of this study is the development of scales for existing constructs, rather than the creation of new constructs, we relied exclusively on the theoretical background, definitions, and interview statements provided by Gerlach and Cenfetelli (2020). Accordingly, we chose to follow a deductive item generation approach, where items are generated based on an existing definition and theoretical knowledge (Hinkin Timothy R., 1998).

First, building on the previous steps, we derived broad categories for both constructs. We refined the initial categories in an iterative manner until reaching sufficiently distinct dimensions. The interview excerpts from Gerlach and Cenfetelli (2020), which are provided in their paper, proved to be particularly insightful. They offered different perspectives on the constructs at hand as perceived by the interviewees.

As a result, for PoA, we derived three areas to categorize issues of not being able to sustain focus on a stimulus. The three areas and their eponymous dimensions are 1) productivity harm 2) physical harm 3) social harm. For SEG, we only determined one dimension, the situational value of information.
Following the recommendation of Lewis et al. (2005), we created multiple item stems for each dimension to particularize their sub-parts. For instance, for PoA's dimension of productivity harm we derived item stems i) not being able to sustain focus on a work-related stimulus ii) not being able to sustain focus on a study-related stimulus, and iii) not being able to sustain focus on a hobby-related stimulus. For PoA's dimension of physical harm we identified the stems i) risking injury and ii) risking material damage. Last, for the dimension of social harm we identified the stems i) risking reputational damage and ii) risking emotional damage. For SEG we derived the item stems i) importance of information and ii) relevance of information to enduring goal as the situational value of information depends on both of these factors (Gerlach and Cenfetelli, 2020). Table 1 provides an overview of the dimensions and item stems for PoA and SEG.

Next, we applied domain sampling, which refers to assigning statements to each item stem and only retaining the ones that best represent an item stem to be the final item statements (Bollen and Lennox, 1991; Lewis et al., 2005). In doing so, we followed the basic principles recommended in item-writing literature to phrase the items as straightforward, simple, and appropriate for the target population as possible, to avoid leading questions as they may induce biases, and not to use items that all respondents will answer similarly (Clark and Watson, 2019; Hinkin Timothy R., 1998). A list of all items can be found in Table 7, which presents the final measurement scales at the end of this paper.

Construct	**Dimension**	**Item Stem**
Problem of Attention (PoA)	Productivity harm	Work-related stimulus
		Study-related stimulus
		Hobby-related stimulus
	Physical harm	Risking injury
		Risking material damage
	Social harm	Risking reputational damage
		Risking emotional damage
Service to Enduring Goal (SEG)	Situational value of information	Importance of information
		Relevance of information to enduring goal

Table 1. Dimensions and item stems for PoA and SEG (In order to report SEG, participants first are to be asked to reflect upon the enduring goal motivating their <app> use).

Moreover, we put careful consideration into developing reflective indicators for PoA and SEG instead of formative ones. While the former implies that the item statements are imperfect reflections of the underlying constructs (Nunally, 1994), the combined items form the construct in the latter (Bollen and Lennox, 1991). Unlike formative indicators, individual reflective indicators can be eliminated in the scale development process without compromising the meaning of the construct as a whole (MacKenzie et al., 2011).

5 Stage II: Instrument

The statements created in Section 4 are used as item statements in the instrument. The instrument is purified and improved through the following two iterations: Pre-test and pilot-test. For each, data is collected from a different set of participants.

5.1 Pre-Test

The pre-test represents the first empirical test for the instrument. It should be conducted among subjects that are knowledgeable about the underlying constructs (Lewis et al., 2005). We identified 21 prospective content experts working as doctorate students or postdocs in German universities. To be deemed content experts, they either had to have experience working on research related to the field, be knowledgeable in construct development theory by Lewis et al. (2005), or both. Ten students and postdocs followed our personalized email invitation and agreed to participate in the pre-test. Detailed instructions on how to complete the instrument were sent to them, asking them to fill out the attached survey for any one app that they frequently use on their phone. We also informed them of the purpose of the pre-test and asked participants for qualitative feedback on the instrument. Specifically, we included a text box at the end of the instrument to collect comments on the “format, content, understandability, terminology, ease and speed of completion” (Lewis et al., 2005, p. 392) as well as their expert judgment on whether or not specific items should be added, deleted or rephrased. Following data collection, we reviewed all comments. Due to the overall very positive feedback, we only had to implement minor changes. Most importantly, the items “My use of <the app> has distracted me while working.” and “My use of <the app> has distracted me while studying.” were combined into one item, as multiple experts commented on the possibility of having participants that have either not worked or studied before.

5.2 Pilot-Test

We administered the adjusted instrument to a small sample of participants similar to the target population for a second empirical test. While the pre-test was sent to experts in a Microsoft word file, the pilot test was distributed as an online questionnaire in a crowdsourcing marketplace. Specifically, we used CloudResearch’s Connect as, compared to other crowdsourcing marketplaces, CloudResearch offers proprietary quality verification features to combat survey fraud and inattentive participants that result in increased data quality (Douglas et al., 2023). Using crowdsourcing marketplaces is a common practice in scientific research (Owens and Bernstein, 2019), and data collected from respondents recruited in this manner has been demonstrated to be externally and internally valid (Berinsky et al., 2012). Items were measured on a 5-point Likert scale. The survey itself was integrated into a proprietary mobile application, “AppUsage”, which was developed by one of the authors. AppUsage is only available for Android devices and is published on the Google PlayStore[1]. Once downloaded, the app asks participants for permission to access usage data stored on their phones. We developed the app in a way to not collect any personal data. Further, within the graphical user interface of the app, participants received an extensive briefing explaining the scope of the collected data. AppUsage communicates with the internal app manager of the Android device to extract data on the frequency and duration of app use, a process called mobile data donations (Ohme et al., 2020). This data is available locally on every Android device to enable services such as the Digital Wellbeing app[2] by Google, presenting the user with summary statistics of their smartphone use. By enabling mobile data donations through AppUsage, we were able to extract use data that reached back 8 days from the moment the app was installed. The aggregated data allowed for the deduction of a participant's most-used apps by the number of times it was launched and by overall use time.

Our newly developed constructs are representative of the intended and unintended consequences of state-tracking habits. Evidently, not every app used on a smartphone is used habitually, and thus not every app represents a state-tracking habit. While apps that are used regularly and are high in use duration can be considered smartphone habits (De Guinea and Markus, 2009; Lally and Gardner, 2013; Polites and Karahanna, 2013; Sigerson et al., 2017), launching apps often throughout the day for several days in a row is an indicator of a state-tracking habit (Gerlach and Cenfetelli, 2020).

[1] https://play.google.com/store/

[2] https://play.google.com/store/apps/details?id=com.google.android.apps.wellbeing&hl=de&gl=US

Thus, we related our constructs to the three most regularly launched apps of each participant[3], instead of simply relating them to a selection of popular apps. In the item statements of PoA and SEG <the app> was automatically replaced by the respective app names within the questionnaire. Therefore, in the pilot-test survey, each participant judged the intended and unintended consequences of the three apps that represent their personal state-tracking habits resulting in a preliminary data set consisting of 120 observations. We included control questions in the survey and eliminated observations in which incorrect answers were given. Moreover, we removed 6 incomplete observations, where apps used by the participant could not be retrieved due to technical problems. Overall, we removed 18 out of the collected 120 observations, yielding a final data set consisting of 102 observations. The goal of the pilot test is to further improve the instrument based on the participant's suggestions. However, we again received strictly positive feedback and no suggestions for further improvements were made. We, therefore, decided to retain the 102 observations for the subsequent exploratory assessment.

5.3 Item screening

The purpose of item screening is to apply a qualitative or quantitative procedure to calculate whether each item adequately represents the content domain of its respective construct (Lewis et al., 2005). As suggested by Lewis et al. (2005), we adapted the Content Validity Ratio (CVR), a quantitative content validity approach developed by Lawshe (1975). This approach stands out due to being quick and easy to perform and not having any disadvantages compared to other content validity methods (Tojib and Sugianto, 2006). First, we selected a panel of six experts knowledgeable about state-tracking habits, who at the time were working as doctorate students or postdocs in German universities. The experts were provided with the updated list of items for both PoA and SEG and asked to evaluate the relevance of each item to its respective construct. The relevance was assessed on a three-point scale with 1 representing 'Not relevant', 2 representing 'Important (But Not Essential)', and 3 representing 'Essential'. Using the resulting data, the content validity ratio (CVR) can be computed using the following formula:

$$CVR = (n - N/2) / (N/2)$$

N represents the total number of experts in the panel and n represents a count of the number of panelists that chose either 3 or 2, deeming the respective item appropriate (Lewis et al., 2005). All items achieved significant content validity ratio as indicated by a CVR value of 0.99 or higher (Ayre and Scally, 2014; Lawshe, 1975), implying some level of content validity for the item (Lewis et al., 2005). Since none of the items achieved non-significant CVRs, no items were dropped from the instruments in the item screening step.

6 Stage III: Measurement Properties

For the sake of method triangulation and to receive stable results, both exploratory and confirmatory assessments of the instrument are conducted with different samples. All factor analyses presented in the following sections were evaluated using R version 3.4.3 (R Core Team, 2022).

6.1 Exploratory assessment

Since no items were excluded in the preceding steps, we continued the analysis with our pilot test data. Following data collection, we computed a response rate of 75%[4], which is much higher than the recommended minimum of 20% (Malhotra and Grover, 1998). When examining statistical power, after

[3] We chose to present each participant with their three most launched apps throughout the eight days prior to their download of AppUsage. We made sure that each of these apps was used every day to preclude that an app was in the top three simply due to being used a lot on one specific day.

[4] Out of 54 participants hired on Amazon Mechanical Turk, 40 finished the task.

data clean-up, we can observe a subject-to-item ratio of 102:13, which again is higher than the recommended minimum of 5:1 (Hair et al., 2018; Lewis et al., 2005). Thus, we judge the response quality to be sufficient for the exploratory assessment.

To assess the factor structure of our constructs, we evaluated both PoA and SEG individually before estimating an exploratory factor analysis with both constructs collectively. For both constructs, the EFA suggested one-factor solutions and factor loadings were reasonably high exceeding the threshold of 0.5 indicating that items are related to one another and thus measure the same construct (Lewis et al., 2005). According to Segars (1998, p. 148) this approach provides "fullest evidence of measurement efficacy". Proceeding with the full model, we checked the Kaiser-Meyer-Olkin (KMO) measure of sampling adequacy and Bartlett's Test of Sphericity to ensure that our data is suitable for factor analysis. The overall KMO value of 0.89 for our constructs is considered highly desirable (Hair et al., 2018). Bartlett's Test of Sphericity delivered the following results: $\chi2(78) = 878.904$, $p = 0.00$. This significant result (<0.05) confirms that our proposed items relate to one another to a sufficient degree to proceed with the EFA. We then applied principal axis factoring with a varimax rotation to extract two factors. We verified that a two-factor solution best suits our data by visualizing the eigenvalues on a scree plot and retaining factors with an eigenvalue above 1 (Lewis et al., 2005; Nunally, 1979).

	Factor 1	**Factor 2**
PoA2	**0.68**	0.06
PoA4	**0.92**	0.05
PoA5	**0.89**	0.02
PoA6	**0.78**	0.00
PoA7	**0.90**	0.10
PoA8	**0.85**	0.06
PoA9	**0.79**	0.14
SEG1	0.04	**0.76**
SEG2	0.02	**0.73**
SEG3	0.21	**0.71**
SEG4	0.00	**0.73**
Variance explained by factor	0.69	0.31

Table 2. Structure matrix.

Literature provides varying recommendations on how high factor loading should be, as low loadings signal that an item is not measuring a construct to a sufficient degree. Although thresholds for acceptable factor loadings of 0.45 and 0.5 have been used in MIS construct development, values greater than 0.5 are preferable for practical significance (Hair et al., 2018). As most of our items loaded on a factor with a significance of 0.68 or above, we chose to remove one item (PoA3) that had a loading of only 0.6. We also eliminated item PoA1 due to cross-loadings on both factors. Squaring each of the loadings and computing their ratio resulted in a value of 1.2989, indicating problematic cross-loading (Hair et al., 2018). Both items were only dropped after careful subjective judgment to ensure that no item with strong theoretical relevance was dropped (Lewis et al., 2005). After dropping items PoA1 and PoA3 we estimated another exploratory factor analysis. Table 2 shows factor loadings for all items of the second iteration of our exploratory factor analysis. Reliability and validity values exceed the recommended thresholds (Hair et al., 2010) and we did not observe any cross-loading.

To assess model fit, we draw on absolute and incremental fit indices. For instance, the Root Mean Square Error of Approximation (RMSEA) rates how well an a priori model reproduces the sample data. Values below 0.06 indicate good fit and values under 0.08 are considered adequate (Hu and Bentler, 1999). The Tucker Lewis Index of factoring reliability (TLI) was 0.974. Values above 0.92 indicate a good fit (Hair et al., 2018). Table 3 summarizes the goodness of fit indices.

Fit index	Acceptable Threshold	Measurement Model
TLI	> 0.92	0.974
RMSEA	< 0.08	0.030

Table 3. Goodness of fit indices of structural model.

Both constructs demonstrate a high internal consistency as demonstrated by Cronbach's Alpha. While high reliability values are desired, values above 0.95 can indicate problems of item redundancy or response patterns (Hair et al., 2018). The Cronbach's Alpha values for PoA (0.94), and SEC (0.82) indicate good reliability without exceeding 0.95.

6.2 Confirmatory assessment

As we dropped two items in the exploratory assessment, we administered the adjusted questionnaire to new participants from an online crowdsourcing marketplace. After dropping 6 incomplete observations, where apps used by the participant could not be retrieved due to technical problems, and 12 observations of participants who failed attention checks, we retained 165 valid and complete observations. We then followed the comprehensive step-by-step guide by Hair et al. (2018) as recommended by Lewis et al. (2005) to assess the efficacy of our measurements. The response rate was 87.14%[5], which, again, is much higher than the recommended minimum of 20% (Malhotra and Grover, 1998).

	Mean	S.D.	Item Loading	Standardized Item Loading	Cronbach's alpha	C.R.	A.V.E.
PoA2	2.880	1.120	1.000	0.744	0.92	0.923	0.641
PoA4	2.640	1.320	1.392	0.840			
PoA5	2.620	1.320	1.298	0.820			
PoA6	2.820	1.260	1.257	0.834			
PoA7	2.890	1.230	1.227	0.833			
PoA8	2.610	1.240	1.139	0.765			
PoA9	2.870	1.090	0.945	0.725			
SEG1	3.450	1.040	1.000	0.771	0.82	0.819	0.531
SEG2	3.500	1.050	0.941	0.717			
SEG3	3.470	1.030	0.913	0.711			
SEG4	3.450	1.050	0.939	0.716			

Table 4. CFA results S.D. = Standard deviation C.R. = Composite reliability A.V.E. = Average variance extracted All standardized item loadings are significant at $p < 0.01$.

We observed a subject-to-item ratio of 165:11, which again is higher than the recommended minimum of 5:1 (Hair et al., 2018; Lewis et al., 2005). First, we ensured normality by plotting the distribution of all variables and examining skewness and kurtosis statistics. These statistics suggested sufficient normality and no excessive kurtosis (Lewis et al., 2005). Next, to rule out an identification problem, we re-estimated our model multiple times with different starting values, as suggested by Lewis et al. (2005). Since results converged at the same point, we assume the absence of an identification problem (Lewis et al., 2005). Table 4 shows the results of the confirmatory factor analysis. All standardized factor loadings are higher than 0.7 and thus indicate convergent validity (Hair et al., 2018). The average

[5] Out of 70 participants hired on Amazon Mechanical Turk, 61 finished the task.

variance extracted (AVE) was greater than 0.5 for both constructs again indicating convergent validity (Hair et al., 2018). Both PoA and SEG feature a composite reliability above the threshold of 0.7 (Hair et al., 2018). Cronbach's Alpha values for PoA (0.92) and SEC (0.82) indicate good reliability.

Again, we confirmed the reliability and validity of our model and assessed how well our model fits the observed data through multiple fit indices. As shown in Table 5, the model fit is similarly high to that of the previous assessment. The value for the Comparative Fit Index (CFI) was 0.95, just at its cutoff value of 0.95 (Hu and Bentler, 1999). The standardized root mean squared residual (SRMR) of 0.058 was below its cutoff value of 0.09. As the commonly used chi-square test is affected by varying sample sizes, we follow the recommendation of considering the ratio of chi-square statistics to degrees of freedom instead (Wheaton et al., 1977). The derived ratio of 2.241 ($\chi2(43) = 96.383$, $p = 0.00$) lies well below its established cutoff value of 3.0 (Hair et al., 2018). After assessing all fit indicators, we analyzed correlation coefficients among all items. Within each construct, all correlation coefficients were statistically significant at $p < 0.01$, indicating that all items are related to one another and thus are measuring the same construct (Lewis et al., 2005). Discriminant validity of PoA and SEG including a related variable analysis will be discussed in the next section.

Fit index	Acceptable threshold	Structural model
$\chi2$ / df	< 3.00	2.241
CFI	> 0.95	0.950
TLI	> 0.92	0.935
RMSEA	< 0.08	0.077
SRMR	< 0.09	0.058

Table 5. Goodness of fit indices of structural model.

6.3 Related variable analysis

Through a review of literature, we derived a list of concepts that, based on their definitions or applications, appear similar to our newly created constructs. In this section, we argue why these concepts are in fact distinct from each other and proceed to prove their discriminant validity empirically.

PoA is distinct from similar concepts such as compulsive smartphone use (Zhang et al., 2014) and habitual smartphone use (Van Deursen et al., 2015) (see Table 6). While both PoA and CSU are concerned with the negative side of ambivalent IS (smartphone) use, PoA captures the degree to which state-tracking habits are responsible for situations in which an individual is unable to sustain focus on a stimulus that requires such focus (Gerlach and Cenfetelli, 2020; Stanko and Beckmann, 2014) and therefore focuses on the undesired outcome of IS use. The measure for CSU on the other hand, including statements such as "I find it difficult to control my <app> use", captures the behavior of compulsively using one's smartphone rather than the resulting consequences of a given behavior. While the two constructs are distinct, we still expect them to be correlated for the following reason: Since compulsive behavior is characterized by "the feeling that one *has to* perform" the behavior, regardless of its appropriateness (Luigjes et al., 2019, p. 10), compulsive smartphone use might result in situations where an individual is unable to sustain focus on a stimulus due to being distracted by its smartphone and thus in problems of attention. Similar to CSU, the measure for habitual smartphone use, including items such as "<the app> is a part of my life" captures a behavior, rather than consequences of a behavior. Further, unlike PoA and CSU, habitual use is not inherently negatively connotated, as habits are often goal-aligned (Ouellette and Wood, 1998; Wood and Neal, 2007; Wood and Rünger, 2016).

Finally, SEG is distinct from the similar concept of perceived usefulness (Davis, 1989). The main difference between the two constructs lies in the ability of SEG to measure the service to enduring goals that are not related to productivity or work. Perceived usefulness, defined as "the degree to which a person believes that using a particular system would enhance his or her job performance" (Davis, 1989, p. 320) features items gauging a system's contributions to job performance, increased productivity, or speed of working. SEG on the other hand allows for individual enduring goals that go beyond the scope of productivity as common enduring goals include maintaining social relationships or maintaining one's identity (Gerlach and Cenfetelli, 2020). While the two constructs are distinct, we expect SEG and PU to be correlated since, depending on the manifestation of an individual's enduring goal, the two constructs can produce similar results.

To empirically examine discriminant validity of PoA and SEG, supplemental data on CSU, HU, and PU, provided by the participants, was included in the administration of the instruments. We adapted all scales to target the use of a single app instead of a smartphone. Next, we estimated another confirmatory factor analysis using the sample containing the supplemental data. All items loaded on their appropriate factor exceeding established cutoff values (Hair et al., 2018) with factor levels ranging from 0.544 to 0.838. As reported in Table 6, both problem of attention and service to enduring goal discriminate from all other related constructs. Correlations between the constructs in the columns and rows were lower than the square root of the average variance extracted at the respective intersection. Further, the observed correlations are in line with our expectations based on the theoretical definitions of the constructs. Problem of attention is highly correlated (0.6967) with compulsive smartphone use but only exhibits low (0.1624) and moderate (0.3291) correlations with habitual use and perceived usefulness. Likewise, service to enduring goal correlates highest (0.6175) with its related construct perceived usefulness. All fit statistics exceeded established conservatory cutoff values (Brown, 2006): CFI was 0.9 and RMSEA was 0.076.

Construct	**Mean**	**S.D.**	**C.R.**	**PoA**	**SEG**	**CSU**	**HU**	**PU**
PoA	2.762	1.016	0.924	**0.8006**				
SEG	3.471	0.838	0.819	0.4721	**0.7287**			
CSU	3.281	0.981	0.828	0.6967	0.4198	**0.7855**		
HU	3.819	0.661	0.799	0.1624	0.6041	0.3776	**0.6458**	
PU	3.714	0.710	0.845	0.3291	0.6175	0.3212	0.5494	**0.6914**

Table 6. Correlation matrix of structural model for related constructs. Bold values along the diagonal of the correlations represent the square root of AVEs.

7 Discussion

This research contributes to research and practice by building on the existing conceptualization of the intended and unintended consequences of state-tracking habits represented by service to enduring goal and problem of attention (Gerlach and Cenfetelli, 2020). We develop scales to measure these concepts empirically, following an established construct development methodology for IS (Lewis et al., 2005). We further demonstrate both constructs' discriminant validity by demonstrating theoretically and empirically that they are distinct from the similar established constructs of compulsive smartphone use (Zhang et al., 2014), habitual smartphone use (Van Deursen et al., 2015), and perceived usefulness (Davis, 1989). We argue that to further our understanding of problematic smartphone use and the theorized preventive role of self-control, being able to quantify the intended and unintended consequences of state-tracking habits is indispensable. With the help of our newly developed constructs, a quantification of the quality of state-tracking habits is now possible. Since the theory of IT-mediated state-tracking was designed to not be restricted by the single-IS perspective (Gerlach and Cenfetelli, 2020, 2021), we gave careful consideration to the applicability of PoA and SEG to various IS use contexts. Table 7 presents the final measurement scales for PoA and SEG. It is important to note that, when deploying the scale for SEG, respondents should be asked to reflect upon the enduring goal motivating the use of the given app before completing the measurement scale.

Construct		Items
Problem of Attention (PoA)	PoA1	My use of <the app> has distracted me while working or studying.
	PoA2	I have risked physically harming myself or others due to being distracted by <the app>.
	PoA3	I have risked causing material damage to something due to my use of <the app>.
	PoA4	I used <the app> in situations where my attention is required, even though it might be dangerous.
	PoA5	My use of <the app> was not perceived well by people in my social environment.
	PoA6	I have hurt other people's feelings by using <the app>.
	PoA7	I used <the app> when I was with my friends or colleagues, even though they might not like it.
Service to Enduring Goal (SEG)	SEG1	Through my use of <the app>, I have obtained information important to my goal.
	SEG2	My use of <the app> has provided me with information crucial to my goal.
	SEG3	The information I have obtained using <the app> is relevant to my goal.
	SEG4	By using <the app>, I obtained information that is in line with my goal.

Table 7. The final measurement scales for PoA and SEG. (The index for PoA has been reset after excluding PoA1 and PoA3).

8 Conclusion, Future Research, and Limitations

By facilitating the evaluation of state-tracking habits based on the extent to which they lead to problems of attention and are of service to one's individual goals, our newly developed scales serve researchers and individuals seeking to reduce their state-tracking alike. Individuals may use the scales to personally reflect on their state-tracking habits, identifying unwanted habits and subsequently applying measures such as self-control strategies precisely for these habits. Similarly, researchers can use our scales to distinguish between state-tracking habits with predominantly intended or unintended consequences, while accounting for differences between apps and individuals, for example when measuring the impact of behavioral interventions. In our sample, the mean PoA associated with any given state-tracking habit was 2.762 and the mean SEG was 3.471. While these values serve as initial starting points for thresholds to identify unproblematic/problematic and useful/useless state-tracking habits, future studies should gather further data and consider the measures' distributions. We further encourage future studies to explore the role of self-control in preventing the execution of state-tracking habits and more importantly the relationship between the consequences of state-tracking habits and an individual's inclination to exert self-control to prevent these habits. Like any empirical study, this study is not free of limitations. As mentioned before, one of the key benefits of the theory of IT-mediated state-tracking is its capability to be applied to many different IS use contexts (Gerlach and Cenfetelli, 2020). In our study, we applied it in the context of smartphone use to motivate the need for, develop, and validate our measures. While state-tracking commonly happens on smartphones, computers, and laptops are also used in state-tracking behaviors as individuals tend to use the digital device they have at hand (Gerlach and Cenfetelli, 2020). Since applying a single IS perspective might misrepresent phenomena and lead to incomplete theories, we put careful consideration into the scales' ability to be adapted to different IS use contexts. Future studies should use these scales to measure consequences of state-tracking habits that occur outside the realm of smartphones to validate the scales' usefulness. To do so, in the wording of the items <the App> can be replaced by any other IS. To conclude, our study contributes to IS research by developing valid and reliable scales for PoA and SEG, enabling further empirical validation of the theoretical model of IT-mediated state-tracking.

Acknowledgements

This work has been partially supported by the DFG as part of the CRC 1053 MAKI.

References

Adriaanse, M. A., Kroese, F. M., Gillebaart, M., and De Ridder, D. T. D. (2014). "Effortless Inhibition: Habit Mediates the Relation between Self-Control and Unhealthy Snack Consumption," *Frontiers in Psychology* (5:MAY), 10–15.

Aranda, J. H., and Baig, S. (2018). "Toward 'JOMO': The Joy of Missing Out and the Freedom of Disconnecting," in *Proceedings of the 20th International Conference on Human-Computer Interaction with Mobile Devices and Services*, Barcelona, Spain, 1–8.

Ayre, C., and Scally, A. J. (2014). "Critical Values for Lawshe's Content Validity Ratio: Revisiting the Original Methods of Calculation," *Measurement and Evaluation in Counseling and Development* (47:1), 79–86.

Berinsky, A. J., Huber, G. A., and Lenz, G. S. (2012). "Evaluating Online Labor Markets for Experimental Research: Amazon.Com's Mechanical Turk," *Political Analysis* (20:3), 351–368.

Bollen, K., and Lennox, R. (1991). "Conventional Wisdom on Measurement: A Structural Equation Perspective," *Psychological Bulletin* (110:2), 305–314.

Boswell, W. R., and Olson-Buchanan, J. B. (2007). "The Use of Communication Technologies after Hours: The Role of Work Attitudes and Work-Life Conflict," *Journal of Management* (33:4), 592–610.

Brown, T. A. (2006). *Confirmatory Factor Analysis for Applied Research*, New York: The Guilford Press.

Chen, C., Zhang, K. Z. K., Gong, X., Zhao, S. J., Lee, M. K. O., and Liang, L. (2017). "Understanding Compulsive Smartphone Use: An Empirical Test of a Flow-Based Model," *International Journal of Information Management* (37:5), Elsevier, 438–454.

Cheng, Y., and Meng, J. (2021). "The Association between Depression and Problematic Smartphone Behaviors through Smartphone Use in a Clinical Sample," *Human Behavior and Emerging Technologies* (3:3), 441–453.

Churchill, G. A. J. (1979). "A Paradigm for Developing Better Measures of Marketing Constructs," *Journal of Marketing Research* (16:1), 64–73.

Clark, L. A., and Watson, D. (2019). "Constructing Validity : New Developments in Creating Objective Measuring Instruments," *Psychological Assessment*, 1–16.

Condliffe, J. (2017). "Constant Phone Checkers Are Totally Strung Out." (https://www.technologyreview.com/2017/02/23/153586/constant-phone-checkers-are-totally-strung-out/).

Davis, F. D. (1989). "Perceived Usefulness, Perceived Ease of Use, and User Acceptance of Information Technology," *MIS Quarterly: Management Information Systems* (13:3), 319–339.

Van Deursen, A. J. A. M., Bolle, C. L., Hegner, S. M., and Kommers, P. A. M. (2015). "Modeling Habitual and Addictive Smartphone Behavior: The Role of Smartphone Usage Types, Emotional Intelligence, Social Stress, Self-Regulation, Age, and Gender," *Computers in Human Behavior* (45), Elsevier Ltd, 411–420.

Douglas, B. D., Ewell, P. J., and Brauer, M. (2023). "Data Quality in Online Human-Subjects Research: Comparisons between MTurk, Prolific, CloudResearch, Qualtrics, and SONA," *Plos One* (18:3), e0279720.

Elhai, J. D., Levine, J. C., O'Brien, K. D., and Armour, C. (2018). "Distress Tolerance and Mindfulness Mediate Relations between Depression and Anxiety Sensitivity with Problematic Smartphone Use," *Computers in Human Behavior* (84), Elsevier Ltd, 477–484.

Fishbein, M., and Ajzen, I. (1980). *Predicting and Understanding Consumer Behavior: Attitude-Behavior Correspondence.*, London: Pearson.

Gardner, B. (2015). "A Review and Analysis of the Use of 'Habit' in Understanding, Predicting and Influencing Health-Related Behaviour," *Health Psychology Review* (9:3), Taylor & Francis, 277–295.

Gerlach, J. P., and Cenfetelli, R. T. (2020). "Constant Checking Is Not an Addiction: A Grounded Theory of IT- Mediated State-Tracking," *MIS Quarterly: Management Information Systems* (44:4), 1705–1731.

Gerlach, J. P., and Cenfetelli, R. T. (2021). "Overcoming the Single-IS Paradigm in Individual-Level IS Research," *Information Systems Research* (March 2022).

Gillebaart, M., and de Ridder, D. T. D. (2015). "Effortless Self-Control: A Novel Perspective on Response Conflict Strategies in Trait Self-Control," *Social and Personality Psychology Compass* (9:2), 88–99.

Gillebaart, M., Schneider, I. K., and De Ridder, D. T. D. (2016). "Effects of Trait Self-Control on Response Conflict About Healthy and Unhealthy Food," *Journal of Personality* (84:6), 789–798.

De Guinea, A. O., and Markus, L. (2009). "Why Break the Habit of a Lifetime? Rethinking the Roles of Intention, Habit, and Emotion in Continuing Information Technology Use," *MIS Quarterly* (33:3), 433–444.

Hair, J. F., Black, W. C., Babin, B. J., and Anderson, R. E. (2018). *Multivariate Data Analysis*, (Eighth Edi.), Cengage Learning EMEA.

Hair, J. F., Black, W. C., Babin, B. Y. A., Anderson, R., and Tatham, R. (2010). "Multivariate Data Analysis," *A Global Perspective*.

Hinkin Timothy R. (1998). "A Brief Tutorial on the Development of Measures for Use in Survey Questionnaires," *Organizational Research Methods* (1:1), 104–121.

Holte, A. J., and Ferraro, F. R. (2020). "True Colors: Grayscale Setting Reduces Screen Time in College Students," *Social Science Journal* (00:00), Routledge, 1–17.

Hu, L. T., and Bentler, P. M. (1999). "Cutoff Criteria for Fit Indexes in Covariance Structure Analysis: Conventional Criteria versus New Alternatives," *Structural Equation Modeling* (6:1), 1–55.

Inzlicht, M., Werner, K. M., Briskin, J. L., and Roberts, B. W. (2021). "Integrating Models of Self-Regulation," *Annual Review of Psychology* (72), 319–345.

Lally, P., and Gardner, B. (2013). "Promoting Habit Formation," *Health Psychology Review* (7:SUPPL1), 37–41.

Lan, Y., Ding, J.-E., Li, W., Li, J., Yifei, Z., Liu, M., and Fu, H. (2018). "A Pilot Study of a Group Mindfulness-Based Cognitive-Behavioral Intervention for Smartphone Addiction among University Students," *Journal of Behavioral Addictions* (7:4), 1171–1176.

Lawshe, C. H. (1975). "A Quantitative Approach To Content Validity," *Personnel Psychology* (28:4), 563–575.

Lee, Y. K., Chang, C. T., Lin, Y., and Cheng, Z. H. (2014). "The Dark Side of Smartphone Usage: Psychological Traits, Compulsive Behavior and Technostress," *Computers in Human Behavior* (31:1), Elsevier Ltd, 373–383.

Lewis, B. R., Templeton, G. F., and Byrd, T. A. (2005). "A Methodology for Construct Development in MIS Research," *European Journal of Information Systems* (14:4), 388–400.

Limayem, M., Hirt, S. G., and Cheung, C. M. K. (2007). "How Habit Limits the Predictive Power of Intention: The Case of Information Systems Continuance," *MIS Quarterly: Management Information Systems* (31:4), 705–737.

Lin, Y. H., Lin, Y. C., Lin, S. H., Lee, Y. H., Lin, P. H., Chiang, C. L., Chang, L. R., Yang, C. C. H., and Kuo, T. B. J. (2017). "To Use or Not to Use? Compulsive Behavior and Its Role in Smartphone Addiction," *Translational Psychiatry* (7:2), Nature Publishing Group, 1–6.

Loid, K., Täht, K., and Rozgonjuk, D. (2020). "Do Pop-up Notifications Regarding Smartphone Use Decrease Screen Time, Phone Checking Behavior, and Self-Reported Problematic Smartphone Use? Evidence from a Two-Month Experimental Study," *Computers in Human Behavior* (102:August 2019), Elsevier, 22–30.

Luigjes, J., Lorenzetti, V., de Haan, S., Youssef, G. J., Murawski, C., Sjoerds, Z., van den Brink, W., Denys, D., Fontenelle, L. F., and Yücel, M. (2019). "Defining Compulsive Behavior," *Neuropsychology Review* (29:1), Neuropsychology Review, 4–13.

MacKenzie, S. B., Podsakoff, P. M., and Podsakoff, N. P. (2011). "Construct Measurement and Validation Procedures in MIS and Behavioral Research: Integrating New and Existing Techniques," *MIS Quarterly: Management Information Systems* (35:2), 293–334.

Maddux, J. E. (1997). "Habit, Health, and Happiness," *Journal of Sport & Excercise Psychology* (19:4), 331–346.

Malhotra, M. K., and Grover, V. (1998). “An Assessment of Survey Research in POM: From Constructs to Theory,” *Journal of Operations Management* (16:4), 407–425.

Meerkerk, G. J., Van Den Eijnden, R. J. J. M., Franken, I. H. A., and Garretsen, H. F. L. (2010). “Is Compulsive Internet Use Related to Sensitivity to Reward and Punishment, and Impulsivity?,” *Computers in Human Behavior* (26:4), Elsevier Ltd, 729–735.

Morahan-Martin, J. (2005). “Internet Abuse: Addiction? Disorder? Symptom? Alternative Explanations?,” *Social Science Computer Review* (23:1), 39–48.

Nunally, J. C. (1979). “Psychometric Theory: Second Edition.,” *Applied Psychological Measurement* (3:2).

Nunally, J. C. (1994). “Psychometric Theory,” *Philosophical Magazine*.

Ohme, J., Araujo, T., Vreese, C. De, and Piotrowski, J. (2020). *Mobile Data Donations : Assessing Self-Report Accuracy , Sample Biases and Predictive Validity of Mobile News Use with the IOS Screen Time Function*, 1–35.

Ouellette, J. A., and Wood, W. (1998). “Habit and Intention in Everyday Life: The Multiple Processes by Which Past Behavior Predicts Future Behavior,” *Psychological Bulletin* (124:1), 54–74.

Owens, J. C., and Bernstein, I. H. (2019). “Using Online Labor Market Participants for Nonprofessional Investor Research: A Comparison of MTurk and Qualtrics Samples.,” *Journal of Information Systems* (33:1), 113–128.

Panova, T., and Carbonell, X. (2018). “Is Smartphone Addiction Really an Addiction?,” *Journal of Behavioral Addictions* (7:2), 252–259.

Polites, G. L., and Karahanna, E. (2013). “The Embeddedness of Information Systems Habits in Organizational and Individual Level Routines: Development and Disruption,” *MIS Quarterly: Management Information Systems*, 221–246.

Qahri-Saremi, H., and Turel, O. (2020). “Ambivalence and Coping Responses in Post-Adoptive Information Systems Use,” *Journal of Management Information Systems* (37:3), 820–848.

Quinn, J. M., Pascoe, A., Wood, W., and Neal, D. T. (2010). “Can’t Control Yourself? Monitor Those Bad Habits,” *Personality and Social Psychology Bulletin* (36:4), 499–511.

R Core Team. (2022). *R: A Language and Environment for Statistical Computing.*, Vienna, Austria: R Foundation for Statistical Computing. (https://www.r-project.org/.).

Regan, T., Harris, B., Van Loon, M., Nanavaty, N., Schueler, J., Engler, S., and Fields, S. A. (2020). “Does Mindfulness Reduce the Effects of Risk Factors for Problematic Smartphone Use? Comparing Frequency of Use versus Self-Reported Addiction,” *Addictive Behaviors* (108), Elsevier, 106435.

Roffarello, A. M., and De Russis, L. (2021). “Understanding , Discovering , and Mitigating Habitual Smartphone Use in Young Adults,” *ACM Trans. Interact. Intell. Syst. 1, 1, Article* (1:1). (https://scholar.google.it/,).

van Rooij, A. J., Schoenmakers, T. M., van de Eijnden, R. J. J. M., and van de Mheen, D. (2010). “Compulsive Internet Use: The Role of Online Gaming and Other Internet Applications,” *Journal of Adolescent Health* (47:1), Elsevier Ltd, 51–57.

Salehan, M., and Negahban, A. (2013). “Social Networking on Smartphones: When Mobile Phones Become Addictive,” *Computers in Human Behavior* (29:6), Elsevier Ltd, 2632–2639.

Segars, A. H. (1998). “Strategic Information Systems Planning: An Investigation of the Construct and Its Measurement,” *MIS Quarterly* (June), 535–546.

Serenko, A., and Turel, O. (2020). “Directing Technology Addiction Research in Information Systems,” *Data Base for Advances in Information Systems* (51:3), 81–96.

Shaffer, H. J., Hall, M. N., and Vander Bilt, J. (2000). “‘Computer Addiction’: A Critical Consideration,” *American Journal of Orthopsychiatry* (70:2), 162–168.

Sigerson, L., Li, A. Y. L., Cheung, M. W. L., and Cheng, C. (2017). “Examining Common Information Technology Addictions and Their Relationships with Non-Technology-Related Addictions,” *Computers in Human Behavior* (75), Elsevier Ltd, 520–526.

Stanko, T. L., and Beckmann, C. M. (2014). “Watching You Watching Me: Boundary Control and Capturing Attention in the Context of Ubiquitous Technology Use,” *Academy of Management Journal*.

Tojib, R. D., and Sugianto, L.-F. (2006). "Content Validity of Instruments in Is Research," *Journal of Information Technology Theory and Application* (8:3), 31–56.

Wang, C., and Lee, M. K. O. (2020). "Why We Cannot Resist Our Smartphones: Investigating Compulsive Use of Mobile Sns from a Stimulus-Response-Reinforcement Perspective," *Journal of the Association for Information Systems* (21:1), 175–200.

Wang, C., Lee, M. K. O., and Hua, Z. (2014). "Understanding and Predicting Compulsive Smartphone Use: An Extension of Reinforcement Sensitivity Approach," *35th International Conference on Information Systems "Building a Better World Through Information Systems", ICIS 2014*, 1–12.

Wheaton, B., Muthen, B., Alwin, D. F., and Summers, G. F. (1977). *Assessing Reliability and Stability in Panel Models*, (8:1977), 84–136.

Wood, W., and Neal, D. T. (2007). "A New Look at Habits and the Habit-Goal Interface," *Psychological Review* (114:4), 843–863.

Wood, W., and Neal, D. T. (2009). "The Habitual Consumer," *Journal of Consumer Psychology* (19:4), Society for Consumer Psychology, 579–592.

Wood, W., and Rünger, D. (2016). "Psychology of Habit," *Annual Review of Psychology* (67:September), 289–314.

Wu, A. M. S., Cheung, V. I., Ku, L., and Hung, E. P. W. (2013). "Psychological Risk Factors of Addiction to Social Networking Sites among Chinese Smartphone Users," *Journal of Behavioral Addictions* (2:3), 160–166.

Zerach, G. (2020). "Emptiness Mediates the Association Between Pathological Narcissism and Problematic Smartphone Use," *Psychiatric Quarterly*, Psychiatric Quarterly.

Zhang, K. Z. K., Chen, C., Zhao, S. J., and Lee, M. K. O. (2014). "Compulsive Smartphone Use: The Roles of Flow, Reinforcement Motives, and Convenience," *35th International Conference on Information Systems "Building a Better World Through Information Systems", ICIS 2014*.