

EFFECTS OF SMARTPHONE CUES AND ONLINE VIGILANCE ON WELL-BEING AND PERFORMANCE

Niklas Johannes



Behavioural
Science
Institute

EFFECTS OF SMARTPHONE CUES AND ONLINE VIGILANCE ON WELL-BEING AND PERFORMANCE

Niklas Johannes

Coverdesign: Ridderprint B.V. | www.ridderprint.nl
Lay-out & Printing: Ridderprint B.V. | www.ridderprint.nl
ISBN: 978-94-6375-707-2

© 2019, Niklas Johannes, Germany

All rights reserved. No part of this thesis may be reproduced, stored in a retrieval system, or transmitted in any form or by any means without prior written permission of the author.

Effects of Smartphone Cues and Online Vigilance on Well-Being and Performance

Proefschrift

ter verkrijging van de graad van doctor
aan de Radboud Universiteit Nijmegen
op gezag van de rector magnificus prof. dr. J.H.J.M. van Krieken,
volgens besluit van het college van decanen
in het openbaar te verdedigen op dinsdag 11 februari 2020
om 14.30 uur precies

door
Niklas Johannes
geboren op 21 oktober 1989
te Mainz, Duitsland

Promotor:

Prof. dr. Moniek Buijzen

Copromotor:

Dr. Harm Veling

Manuscriptcommissie:

Prof. dr. Sabine Geurts (voorzitter)

Prof. dr. Marjolijn Antheunis (Tilburg Universiteit)

Dr. Tilo Hartmann (Vrije Universiteit Amsterdam)

Effects of Smartphone Cues and Online Vigilance on Well-Being and Performance

Doctoral Thesis

to obtain the degree of doctor
from Radboud University Nijmegen
on the authority of Rector Magnificus Prof. Dr. J. H. J. M. Krieken,
according to the decision of the Council of Deans,
to be defended in public on Tuesday, 11th of February 2020,
at 14:30

by
Niklas Johannes
born on October 21, 1989
in Mainz, Germany

Promoter:

Prof. Dr. Moniek Buijzen

Co-promoter:

Dr. Harm Veling

Manuscript Committee:

Prof. Dr. Sabine Geurts (Chair)

Prof. Dr. Marjolijn Antheunis (Tilburg University)

Dr. Tilo Hartmann (Vrije Universiteit Amsterdam)

TABLE OF CONTENTS

Chapter 1	Introduction	9
Chapter 2	Mind-Wandering and Mindfulness as Mediators of the Relationship Between Online Vigilance and Well-Being	23
Chapter 3	The Relationship Between Online Vigilance and Affective Well-Being in Everyday Life: Combining Smartphone Logging with Experience Sampling	35
Chapter 4	Hard to Resist? The Effect of Smartphone Visibility and Notifications on Response Inhibition	55
Chapter 5	No Evidence that Smartphone Notifications Lead to Goal-Neglect	73
Chapter 6	Social Smartphone Apps Do Not Capture Attention Despite Their Perceived High Reward Value	85
Chapter 7	Simple Motor Responses Change Behavior Through Changes in Explicit Liking: Influencing Preferences for Smartphone Apps	105
Chapter 8	Discussion	131
	References	147
	Samenvatting	171
	Summary	177
	Acknowledgements	183
	About the Author	187

Introduction



“Every once in a while, a revolutionary product comes along that changes everything”, announced former CEO of Apple, Steve Jobs, on January 9th, 2007. He was talking about the first true smartphone, the iPhone. In the following years, the iPhone took the world by storm and other manufacturers followed suit. Today, smartphones are everywhere (CBS, 2018; Pew Research Center, 2017). Smartphones were supposed to be the great democratizer, the ultimate productivity tool, the always-present entertainment machine. However, they also stirred up concerns about users tethered to their devices (Carr, 2011; Turkle, 2012). These concerns were not new. In fact, concerns have surrounded every technological innovation that rapidly gained popularity. When television became mainstream, many public voices expressed unease about a culture lost in entertainment, unable to concentrate (Postman, 1985). But televisions are technologically far inferior to smartphones. In fact, smartphones *are* televisions – and much more. There are apps for almost any purpose, contributing to worries about constantly distracted users. So, are concerns about effects of smartphones on concentration and well-being justified this time around? In this dissertation, we address these concerns.

It looks like even the industry thinks smartphone use, particularly app use, can be problematic for users. More than ten years after the iPhone was introduced, both Apple and Google rolled out a feature that allowed users to see for how long they use their apps, including an option to restrict access to certain apps. These features were the companies’ answer to many users complaining about being “permanently online and permanently connected” (Klimmt, Hefner, Reinecke, Rieger, & Vorderer, 2018; Vorderer & Kohring, 2013). Specifically, people report an increasing apprehension that they cannot refrain from using their phones despite wanting to focus on other, more important tasks (Calderwood, Ackerman, & Conklin, 2014; Näsi & Koivusilta, 2013; Panek, 2014; Pew Research Center, 2018). These concerns may be warranted, as smartphone use that interrupts other tasks has shown to be costly for performance (Chein, Wilmer, & Sherman, 2017; Chein et al., 2017; Q. Chen & Yan, 2016; Kushlev & Dunn, 2015; Kushlev, Hunter, Proulx, Pressman, & Dunn, 2019; Wilmer & Chein, 2016). It seems that smartphones present a source of distraction and people experience conflict when tempted to use their phones instead of performing other tasks (Hofmann, Vohs, & Baumeister, 2012).

Interestingly, recent investigations suggest that receiving a notification (Kushlev, Proulx, & Dunn, 2016; Stothart, Mitchum, & Yehnert, 2015) or even the mere presence of a smartphone (Thornton, Faires, Robbins, & Rollins, 2014; Ward, Duke, Gneezy, & Bos, 2017) may be sufficient to distract from a main task. These findings suggest that people learn that smartphones, and smartphone apps, connect users to their social network, provide entertainment, and grant access to information (Antheunis, Vanden Abeele, & Kanters, 2015; Karapanos, Teixeira, & Gouveia, 2016; Tanis, Beukeboom, Hartmann, & Vermeulen, 2015; van Koningsbruggen, Hartmann, Eden, & Veling, 2017). In turn, *smartphone cues* such as app icons, notifications, and the mere visual presence of a smartphone may attract attention. According to such a view, both the mere presence of smartphones and visual exposure to smartphone app icons can serve as cues that remind users of the connection

apps afford (Carolus et al., 2018; Kardos, Unoka, Pléh, & Soltész, 2018; Ward et al., 2017). Recent work suggests that repeated use, paired with constant exposure to smartphone cues can lead to a psychological state termed *online vigilance* (Klimmt et al., 2018; Reinecke et al., 2018). Users high in online vigilance display a constant awareness of online streams of communication. This vigilance becomes particularly pronounced when people encounter smartphone cues. Therefore, exposure to smartphone cues might come at the cost of dividing attention between the online sphere and a current task (Baumgartner, van der Schuur, Lemmens, & te Poel, 2017; Birnholtz, Davison, & Li, 2017).

Consequently, it stands to debate whether exposure to smartphone cues and the resulting online vigilance affect basic cognitive functions. Equally important, users often complain about how bothersome they find being in a mindset of constant connectivity (e.g., Mihailidis, 2014), hinting at possible negative effects on well-being. Smartphone ownership is rapidly approaching saturation in the Western world (CBS, 2018; Pew Research Center, 2017), making constant connectedness the norm. However, there is little research systematically examining the effect of smartphone cues and online vigilance. Moreover, if smartphone cues indeed have negative effects, there is a need for research testing ways to reduce the appeal of smartphone cues. In this dissertation, we address this gap in the literature. We had the overall aim to investigate *the effects of smartphone cues and online vigilance on well-being and performance*. We approached this general aim with two specific goals. First, we examined whether and how online vigilance is related to people's well-being. Second, we investigated whether smartphone cues influence people's performance. Based on our findings, we formulated a third and fourth goal: Third, we examined why smartphone cues are perceived as distracting. Fourth, we tested whether we can reduce the appeal of smartphone cues for people.

Chapters 2 and 3 address the first goal, testing whether and how online vigilance is related to well-being on the trait level (Chapter 2) and on the state level in everyday life (Chapter 3). Addressing the second goal, in Chapters 4 and 5, we test whether smartphone cues (i.e., visual exposure to smartphones and smartphone notifications) induce online vigilance and interfere with basic cognitive control functions. Addressing the third goal, in Chapter 6 we examine whether the distracting effects of smartphone cues can be explained by the rewarding nature of smartphone cues. Finally, addressing the fourth goal, in Chapter 7 we investigate whether we can reduce people's preferences for specific smartphone cues (i.e., the appeal of smartphone cues). All chapters are in the form of a journal article and are published or submitted for publication. For an overview of the research goals and corresponding chapters, see Table 1.

Table 1 Overview of research goals and chapters

	Research Goals	Chapter 2	Chapter 3	Chapter 4	Chapter 5	Chapter 6	Chapter 7
Goal 1	Does Online Vigilance Relate to Well-Being?	X	X				
Goal 2	Do Smartphone Cues Impact Performance?			X	X		
Goal 3	Are Smartphone Cues Rewarding?					X	
Goal 4	Can We Reduce Preferences for Smartphone Cues?						X

GOAL 1: DOES ONLINE VIGILANCE RELATE TO WELL-BEING?

Previous research suggests that smartphones enable people to connect to others and satisfy social and informational needs at all times via apps (Bayer, Ellison, Schoenebeck, Brady, & Falk, 2018; Jung & Sundar, 2018; Karapanos et al., 2016; van Koningsbruggen et al., 2017). Because users learn to associate smartphone notifications with need gratification, they become sensitive to smartphone cues. For instance, people respond to app notifications almost instantly, even if their phone is in silent mode (Pielot, Church, & de Oliveira, 2014; Pielot, de Oliveira, Kwak, & Oliver, 2014). In addition to responding to notifications, users also initiate smartphone use by themselves to obtain need gratification (Jung & Sundar, 2018; Sundar & Limperos, 2013). Consequently, there is evidence that even the sight of their smartphone may serve as a cue for the learned associations between smartphones and need gratification (Kardos et al., 2018; Ward et al., 2017). Taken together, both the presence of a smartphone and smartphone notifications can be regarded as smartphone cues.

Recently, Reinecke et al. (2018) proposed that these technological cues lead to psychological connectedness to the online world. This connectedness may turn into online vigilance. In the psychological literature, vigilance is defined as the “the ability of organisms to maintain their focus of attention and to remain alert to stimuli over prolonged periods of time” (Warm, Parasuraman, & Matthews, 2008, p. 433). Vigilance in the psychological tradition is demanding, because people need to concentrate and remain alert for long periods of time, resulting in fatigue (Warm, Finomore, Vidulich, & Funke, 2015). Analogously, users high in online vigilance are constantly aware of online communication and alert to respond to smartphone cues (Klimmt et al., 2018; Reinecke et al., 2018). Online vigilance is expressed in three dimensions. Salience refers to mental preoccupation with the online sphere, that is, thoughts about online interactions. Reactibility refers to the responsiveness to online cues, that is, how sensitive users are to smartphone notifications and how quickly

they respond to them. Monitoring refers to a motivation to stay connected, that is, how frequently users observe and check in to their online sphere.

Just like people experience vigilance in a traditional sense as taxing, there is evidence that they can experience online vigilance as bothersome (Mihailidis, 2014; Näsi & Koivusilta, 2013; Pew Research Center, 2018). While online vigilance also has the potential for positive effects on well-being by reminding people of their social circle, online vigilance appears most detrimental when it presents a form of distraction and absentmindedness (Reinecke, 2018). Permanent responsiveness might manifest in increased cognitive load, which can be fatiguing and is related to increased stress (Reinecke, Aufenanger, et al., 2017; Warm, Matthews, & Finomore Jr, 2018). Constant awareness of the online sphere might also come at the expense of focusing on a task or enjoying the present moment (Reinecke, 2018). Although people report to be bothered by a constant alertness of the online sphere, there is little evidence to investigate such a claim. The only study investigating online vigilance found a relation between online vigilance and perceived stress (Reinecke et al., 2018), lending credence to the reported perception of online vigilance as bothersome. However, this does not answer the question whether concerns over online vigilance are warranted. In other words, the question remains whether online vigilance is related to decreased well-being. We conducted two studies to test this idea, reported in Chapters 2 and 3.

In Chapter 2, to obtain a first insight into how online vigilance relates to well-being, we report a survey with 371 participants (Johannes et al., 2018). For the study, we proposed a theoretical model with absentmindedness as possible mechanism to explain why people experience online vigilance as negative. Importantly, we followed recent theoretical advances on well-being and considered both aspects of well-being (Diener, Lucas, & Oishi, 2018): participants reported both the affective component of their well-being (how they usually feel), and the cognitive component of their well-being (how they evaluate their lives). To test the proposition that online vigilance can induce absentmindedness, we relied on two established concepts representing absentmindedness: mind-wandering and mindfulness. We predicted that online vigilance would negatively relate to both well-being outcomes not directly, but indirectly through increased mind-wandering and decreased mindfulness.

Indeed, we found that online vigilance was indirectly related to well-being, but only through decreased mindfulness; mind-wandering played a negligible role. This indirect effect was small, raising doubts whether concerns about the effects of online vigilance are warranted. In addition, a cross-sectional survey leaves many questions unanswered. Specifically, it is unclear whether such a small effect on the trait level will have consequences for well-being and behavior in everyday life. For instance, in the domain of self-control, self-reported trait self-control was not related to behavioral measures of inhibition (Saunders, Milyavskaya, Etz, Randles, & Inzlicht, 2018). Whereas the survey gave us a first impression that online vigilance is weakly related to well-being, we wanted to investigate whether we could find this relation in people's everyday lives.

In Chapter 3, to gain insight into online vigilance in a more ecological setting, we conducted a study combining smartphone logging with experience sampling with 75 participants. Testing the association of online vigilance with well-being in the field presents an important step that addresses many problems of our cross-sectional study (Andrews, Ellis, Shaw, & Piwek, 2015; Miller, 2012). For one, it allowed us to measure online vigilance with behavioral indicators by logging smartphone use. Estimating media behavior is notoriously difficult; hence, recent work recommends to employ objective measures (Ellis, 2019; Ellis, Davidson, Shaw, & Geyer, 2019; Vanden Abeele, Beullens, & Roe, 2013). Second, we could obtain measures of well-being in the moment. Both steps raised ecological validity and enabled us to test the relation between online vigilance and well-being in people's actual lives.

For this study, we aimed to translate the three dimensions of online vigilance to self-reports in the moment and to objective behavioral indicators where possible. We followed the theoretical arguments that online vigilance develops as a consequence of the social features of smartphones (Reinecke et al., 2018). Hence, in our measurements, we focused on social interactions via smartphones. Participants indicated their salience by reporting how frequently they had been thinking about online communication. Furthermore, we followed recent advances in research on daydreaming which shows that the valence of such thoughts plays a key role: Thoughts distracting from the current moment result in decreased well-being, whereas thoughts about friends and family can increase well-being (Franklin et al., 2013; Poerio, Totterdell, Emerson, & Miles, 2016). Thus, we also measured whether thoughts about online communication were positive or negative. As for the behavioral measures, we assessed monitoring as how much time participants spent on social apps before answering a survey. We assessed reactivity as how quickly participants responded to survey notifications.

Overall, the relation between online vigilance and well-being in everyday life was negligible. We found that thoughts about online communication were followed by slightly worse well-being. It was much more important, though, whether these thoughts were positive or negative. Contrary to the perceptions of many users (e.g., Näsi & Koivusilta, 2013), other indicators of online vigilance displayed a negligible association with well-being. Furthermore, there are important limitations to our logging study. The design did not allow us to test for causality. Because it is difficult to manipulate well-being, we relied on observational data. Hence, the logging study only allows us to conclude small relations between online vigilance and well-being. However, this relation does not inform us whether smartphone cues have an effect on performance, another prominent concern in the public debate (Carr, 2011; Turkle, 2012). In fact, performance can be manipulated in experimental designs. Therefore, in Chapters 4 and 5 we focused on the second goal of this dissertation, investigating whether smartphone cues have an effect on online vigilance and performance.

GOAL 2: DO SMARTPHONE CUES IMPACT PERFORMANCE?

1

Many smartphone users complain that constant exposure to their smartphones and smartphone notifications distract them from their main tasks (Mihailidis, 2014; Näsi & Koivusilta, 2013). In other words, smartphone cues might impair performance. Such an assumption is in line with several theories that assume that smartphone cues signal rewarding experiences to users (Bayer, Campbell, & Ling, 2015; Bayer & LaRose, 2018), because smartphones can gratify social needs and needs for entertainment (Du, Kerkhof, & van Koningsbruggen, 2019; Jung & Sundar, 2018; Karapanos et al., 2016; van Koningsbruggen et al., 2017). Consequently, smartphone cues may be constant distractions, competing with more important tasks for students' attention (Hofmann, Vohs, et al., 2012). Accumulating evidence in the field of media multitasking shows that using smartphones or responding to smartphone cues during a task does not only lead to increased time needed to perform a task (Bowman, Levine, Waite, & Gendron, 2010); multitasking also relates negatively to academic performance, both in cross-sectional (Baumgartner et al., 2017; Q. Chen & Yan, 2016; Jeong & Hwang, 2016; Parry & le Roux, 2019) and in experimental studies (Chein et al., 2017; Dietz & Henrich, 2014; Smith, Isaak, Senette, & Abadie, 2011).

Many authors suspect media multitasking affects performance in the form of basic cognitive control functions (van der Schuur, Baumgartner, Sumter, & Valkenburg, 2015): When switching from one task to another, people disengage from the primary task and engage with the secondary task, resulting in so-called switch costs, leading to decreased performance (Strobach, Liepelt, Schubert, & Kiesel, 2012). In other words, by dividing attention between their smartphones and a task, people experience performance decrements. According to this account, users get accustomed to constantly dividing their attention, effectively deploying cognitive control to process several stimuli at once. In essence, this line of reasoning assumes that the constant temptations of media affect basic cognitive functions, so-called executive control. Executive control involves "a set of general-purpose control mechanisms, often linked to the pre-frontal cortex of the brain, that regulate the dynamics of human cognition and action" (Miyake & Friedman, 2012, p. 8).

We were interested whether such temptations in the form of smartphone cues indeed draw on executive control. Previous work supports the idea that interaction is not necessary to impair task performance. Shelton et al. (2009) provided evidence that hearing a phone ring during a task not only resulted in diminished performance; it also led to increased recovery time from the distraction compared to random tones. Supporting their findings, Stothart et al. (2015) conducted an innovative experiment, demonstrating that receiving a notification (without actually checking the notification) led to more errors in a sustained attention to response task. In such a view, smartphones elicited vigilance, which interfered with keeping the task goal in mind, leading to goal-neglect (i.e., decreased performance). Moreover, the authors concluded that the magnitude of the effect was

comparable to that of actual phone interaction. Thornton et al. (2014) employed a similar design; in two experiments, they found the mere presence of a phone to be detrimental to task performance with high cognitive demands. These findings were replicated in a larger sample, providing evidence that online vigilance might tax executive control (Ward et al., 2017).

These studies suggest that smartphone cues (i.e., the mere presence of a smartphone and notifications) affect task performance. Following the proposition made by previous research (van der Schuur et al., 2015), we hypothesized that smartphone cues trigger responses and thoughts that impair various executive control functions (Diamond, 2014; Miyake & Friedman, 2012). Such a mechanism could also explain why users report performance decrements because of a mindset of connectivity: Smartphone cues trigger greater online vigilance, which interferes with executive control. We set out to test the effect of the mere presence of smartphones and smartphone notifications on two crucial executive control functions that were not previously examined and that are important to perform well: inhibition and working memory.

In Chapter 4, we conducted an experiment as a registered report, which means the paper was peer-reviewed both before and after data collection (Chambers, Dienes, McIntosh, Rotshtein, & Willmes, 2015; Nosek & Lakens, 2014). In the experiment, we investigated the effect of smartphone cues on inhibition with 154 students (Johannes, Veling, Verwijmeren, & Buijzen, in press). We manipulated smartphone visibility and notifications. In one group, participants had their phone next to them on the table, whilst doing a stop-signal task, a task commonly used to measure inhibition (Logan, 1994; Verbruggen et al., 2019). In a second group, participants received three notifications before each block of the task, but were not allowed to check these notifications. We expected that these smartphone cues would induce greater online vigilance. Participants may need to inhibit reaching for their phones as well as inhibit thoughts about the smartphone cue. We predicted that exerting this executive control means participants would perform worse on a simultaneous inhibition task. In addition, recent work on inhibition argues that the stop-signal task, our measure of inhibition, not only measures inhibition, but also other attentional processes (Verbruggen, Aron, Stevens, & Chambers, 2010; Verbruggen, Stevens, & Chambers, 2014). Hence, we employed a version of the stop-signal task that enabled us to distinguish between inhibition and other, non-inhibitory processes. This way, we obtained a more valid measure of inhibition whilst exploring whether smartphone visibility and notifications influenced non-inhibitory processes.

Contrary to our predictions, we found no difference in inhibition between the two smartphone conditions and a control condition. To be exact, we found anecdotal evidence for a lack of an effect, both for inhibition and other processes. Interestingly, there was a mismatch between what participants experienced and their actual performance. Participants reported to experience high online vigilance when their smartphone was on the table or received a notification; yet, this online vigilance did not translate to effects on task performance. These findings were surprising to us, and we set out to investigate

whether the lack of an effect applied to working memory, a second crucial executive control function needed for performance.

In Chapter 5, we conducted a second experiment with 39 participants¹ (Johannes, Veling, & Buijzen, 2019). This experiment was designed to test our assumption that smartphone cues induce task-irrelevant thoughts. That is, we tested whether smartphone notifications would lead to goal-neglect. We reasoned that smartphone notifications would impair working memory, as receiving a notification triggers communication goals (e.g., connecting to others). We predicted that these goals would compete with other goals in working memory (e.g., a primary task). To measure this process, we employed a version of the Stroop task that allowed us to assess the extent to which participants would neglect a goal in working memory (Kane & Engle, 2003). Again, one group received three notifications at the beginning of each block; the control group did not receive notifications and had their phone in silent mode. If smartphone cues indeed trigger online vigilance and draw on working memory, we should observe an increase in goal-neglect for participants who received notifications.

Just like in Chapter 4, there was evidence in the data that the manipulation did not affect executive control. Participants in the notification condition did not report more goal-neglect than participants in the control condition. Again, participants reported to be highly distracted by the notifications, yet this distraction did not manifest in behavior². In fact, there was even weak evidence that those in the notification condition performed better. Consistent with our previous findings, there appears to be a mismatch between people's perception of how distracting smartphone cues are and their actual impact. Hence, we wanted to know how people develop such a perception of smartphone cues as distracting and possibly detrimental, addressing our third goal.

GOAL 3: ARE SMARTPHONE CUES REWARDING?

Although smartphone cues and the resulting online vigilance do not appear to pose a problem for well-being or performance in our studies, people still perceived them as distracting. This raises the question of why smartphone cues are perceived as distracting. Most theories assume that smartphones are such salient distractors because they can gratify social needs (Bayer et al., 2015; Reinecke et al., 2018). They propose that people have a strong need for connectedness (Baumeister & Leary, 1995; Deci & Ryan, 2000) and that people can meet these needs via the many social smartphone apps (Jung & Sundar, 2018; Karapanos et al., 2016; Reich, Schneider, & Heling, 2018). In such a view, people repeatedly gratify such social needs, thereby starting to associate smartphone cues with social reward.

¹ The sample size is significantly lower compared to Chapter 4 for two reasons: First, we had one condition fewer. Second, we reached our preregistered stopping rule earlier as a result of our Bayesian sequential sampling method.

² Chapter 5 was conducted before Chapter 4, which is why we only measure distraction here rather than online vigilance.

Such a view is in line with basic human behavior. People generally seek out rewards (Braver et al., 2014), thus giving attentional priority to rewarding stimuli (Anderson, 2016b, 2016a; Anderson, Laurent, & Yantis, 2011a; Chelazzi, Perlato, Santandrea, & Della Libera, 2013; Le Pelley, Mitchell, Beesley, George, & Wills, 2016). In other words, people are not only driven by stimuli in the environment, for instance, the buzzing of a notification. They are also guided by their current motivational states, such as seeking a rewarding experience (Botvinick & Braver, 2015), for instance, the social connection they expect from a notification. People may have learned to associate smartphone cues with rewarding experiences, which, in turn, guide their attention and behavior (Le Pelley et al., 2016; Pool, Brosch, Delplanque, & Sander, 2016). If indeed smartphone cues represent reward to users, this reward could also explain the emergence of online vigilance.

In Chapter 6, we set out to test this assumption (Johannes, Dora, & Rusz, 2019). We wanted to know if prominent social smartphone apps are perceived as rewarding. To that end, we conducted an experiment with 117 participants employing a commonly used task to assess reward-driven attention (Anderson et al., 2011a). During this task, participants had to identify a target, while being distracted by smartphone cues (i.e., different app symbols). In addition, half of the participants handed in their phone an hour before the task. Deprivation of rewarding experiences often strengthens their appeal (Seibt, Häfner, & Deutsch, 2007), leading us to expect that app symbols would be particularly distracting in this group. In addition, to contrast behavior with participants' perception, we conducted an online study with 158 participants to ask people explicitly how rewarding they found different apps.

Once again, results did not support our predictions. Social smartphone apps high in social reward did not attract attention compared to a set of non-social smartphone apps, and compared to apps low in social reward. This lack of a behavioral effect stood in contrast to self-reports of participants. People reported high reward for social apps, but low rewards for non-social apps. In line with our findings regarding smartphone cues and performance, people perceived social apps as rewarding, but this reward could not be detected behaviorally. This consolidates the impression that there is a mismatch in people's perception of the actual impact of mobile technology.

Nonetheless, even if there is a disconnect between perception and behavioral impact, people still complain about phones and their impact. Hence, as a last step we wanted to know whether we could reduce liking and preferences for smartphone cues, addressing our fourth and last goal.

GOAL 4: CAN WE REDUCE PREFERENCES FOR SMARTPHONE CUES?

1

In Chapter 6 we could not detect social reward of smartphone cues in the form of app icons. Yet people still perceive apps as rewarding. Thus, decreasing the explicit perception of apps as appealing might still present an avenue to reduce distraction by smartphone cues. To this end, we employed a so-called go/no-go training (Veling, Lawrence, Chen, van Koningsbruggen, & Holland, 2017). During go/no-go training, participants execute simple motor responses to images of some stimuli (go stimuli) and withhold a response to images of other stimuli (no-go stimuli). Participants subsequently show lower liking for no-go stimuli (Z. Chen, Veling, Dijksterhuis, & Holland, 2016; Clancy, Fiacconi, & Fenske, 2019; Driscoll, de Launay, & Fenske, 2018; Scholten, Granic, Chen, Veling, & Luijten, 2019). More importantly, when food stimuli are used, no-go stimuli are also chosen less for consumption (Allom, Mullan, & Hagger, 2016; Jones et al., 2016), even weeks after the training, indicating reduced preferences for no-go stimuli (Z. Chen, Holland, Quandt, Dijksterhuis, & Veling, 2019).

Therefore, the go/no-go training promises to be an effective intervention tool to decrease liking and preferences for smartphone cues in the form of apps. In Chapter 7 we report two experiments for which we recruited 150 iPhone users, following the standard procedure for the training (e.g., Z. Chen et al., 2016). Before each experiment, participants locked their phones away for an hour in order to ensure apps were perceived as at least somewhat attractive. In the first experiment, participants first rated how much they liked a variety of apps. The most liked apps were then selected for the training. Afterwards, participants rated their liking of these apps once more. In the second experiment, participants did the same procedure, but came back the next day to do a choice task where they chose between go and no-go apps. We predicted that the training would lead to less liking for no-go apps and that participants would choose no-go apps less. We also expected that this liking would mediate the effect of the training on choice.

Both experiments supported our predictions. Indeed, when people did not respond to smartphone apps during go/no-go training, they liked those apps less compared to apps they responded to. Likewise, participants chose such apps less for actual use. Crucially, liking appears to play a major role in the effect of the training on behavior change, as it partially mediated the effect of the training on choice: The training made people like certain apps less, which was partly responsible for people choosing those apps less. As a consequence, despite its simplicity, the go/no-go training presents a promising intervention tool to modify people's responses to smartphone cues.

A WORD ON OPEN SCIENCE AND DATA AVAILABILITY

The Social Sciences currently find themselves in a crisis, as many prominent findings do not replicate (Camerer et al., 2018; Open Science Collaboration, 2015). This lack of robustness is likely due to flexible data analyses (Gelman & Loken, 2013; Simmons, Nelson, & Simonsohn, 2011) and questionable research practices (John, Loewenstein, & Prelec, 2012; Vermeulen & Hartmann, 2015). To restrict so-called researcher degrees of freedom, we followed current best practices and preregistered all studies (Munafò et al., 2017; Nelson & Simmons, 2018; Wagenmakers, Wetzels, Borsboom, van der Maas, & Kievit, 2012). In addition, we made all materials, data, and analysis scripts of published chapters publicly available on my Open Science Framework profile (<https://osf.io/j5xh8/>). For chapters under review, we provide view-only links for all data and materials in the Method sections of the respective chapters. Data management followed the Research Data Management protocol of the Behavioural Science Institute. Accordingly, we registered all published chapters and their corresponding data sets on the Research Information Services of the Radboud Repository.

Mind-Wandering and Mindfulness as Mediators of the Relationship Between Online Vigilance and Well-Being

This chapter has been published as 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>



ABSTRACT

As mobile technology allows users to be online anywhere and at all times, a growing number of users report feeling constantly alert and preoccupied with online streams of online information and communication—a phenomenon that has recently been termed online vigilance. Despite its growing prevalence, the consequences of this constant orientation towards online streams of information and communication for users' well-being are largely unclear. In the present study, we investigated whether being constantly vigilant is related to cognitive consequences in the form of increased mind-wandering and decreased mindfulness, and examined the resulting implications for well-being. To test our assumptions, we estimated a path model based on survey data ($N = 371$). The model supported the majority of our preregistered hypotheses: Online vigilance was indeed related to mind-wandering and mindfulness, but only mindfulness mediated the relation with decreased well-being. Thus, those mentally preoccupied with online communication were overall less satisfied with their lives and reported less affective well-being when they also experienced reduced mindfulness.

Mobile technology, especially smartphones, have become a central part of people's lives (Cumiskey & Ling, 2015) and afford users to be constantly connected to online streams of communication and interaction (Bayer et al., 2015; Mascheroni & Vincent, 2016; Vorderer & Kohring, 2013). In other words, users are "permanently connected and permanently online"⁴. Interestingly, many users complain about the challenges of being in a constant *mindset* of connectivity (Johannes et al., in press; Mihailidis, 2014; Näsi & Koivusilta, 2013). This mindset has recently been defined as *online vigilance*, a state of constant awareness of ongoing threads of online communication and interaction (Klimmt et al., 2018). However, there is little research on the possible consequences of this new mindset for well-being. As absentmindedness has shown to decrease well-being (Friese & Hofmann, 2016; Killingsworth & Gilbert, 2010), a constant division of attention between the present situation and past, ongoing, or future online interactions may result in the same effect. More specifically, this constant division could come at the cost of decreased attentional focus (Thomson, Besner, & Smilek, 2015), which in turn decreases well-being. With the current study, we therefore test whether online vigilance is negatively related to well-being through increased mind-wandering and decreased mindfulness.

ONLINE VIGILANCE AND WELL-BEING

Online vigilance refers to a mindset of constant awareness of online communication and comprises three dimensions (Klimmt et al., 2018). First, *salience* refers to thoughts about past, present, or future online interactions, that is, the intensity and permanence of a mental preoccupation with online streams of information. Second, *reactibility* refers to how responsive a user is to incoming smartphone stimuli, that is, the sensitivity to notifications and speed with which they are checked. Third, *monitoring* refers to how frequently a user checks her or his mobile device, that is, the continuous observation of ongoing threads of online interaction, unprompted by incoming notifications. Online vigilance is markedly nonpathological; whereas problematic smartphone or Internet use are, by definition, maladaptive (Marino, Gini, Vieno, & Spada, 2018), online vigilance describes an acquired mindset that can be both adaptive and maladaptive.

In particular, as Reinecke (2018) lays out, online vigilance bears the potential to foster but also hamper well-being. That is, there are different mechanisms that can account for adaptive, but also maladaptive effects of online vigilance. On the one hand, online vigilance can take the form of awareness of one's social network and social support. In addition, constant access to pleasant content, distractions from unpleasant experiences, and gratification of social needs can be beneficial (Karapanos et al., 2016; Mascheroni & Vincent, 2016; van Koningsbruggen et al., 2017). Therefore, online vigilance might positively contribute to well-being. On the other hand, constantly monitoring and checking online streams of information can induce absentmindedness and possibly distract from a pleasant moment (Shin & Shin, 2016), resulting in decreased well-being. Thus, the link between online vigilance and well-being likely follows different mechanisms, allowing for

both positive and negative effects (Reinecke & Hofmann, 2016; Valkenburg & Peter, 2013). Consequently, rather than assuming a direct relationship, we examined, and preregistered, one possible mediating mechanism, proposing that online vigilance is related to decreased well-being through increased absentmindedness.

MIND-WANDERING AND MINDFULNESS

In order to investigate the notion that online vigilance would be related to absentmindedness, we selected two traits that have been well-researched and present excellent measures to approach the phenomenon of absentmindedness: mind-wandering and mindfulness. Whereas mind-wandering in the form of task-unrelated thoughts can be understood as a general form of absentmindedness (Mooneyham & Schooler, 2013; Smallwood & Schooler, 2015), mindfulness is considered the ability to focus attention fully on the present moment without letting attention wander off, while simultaneously taking a nonjudgmental stance toward thoughts about the present moment (Brown & Ryan, 2003; Sauer et al., 2013). Moreover, both concepts have been linked such that mind-wandering occurs when mindfulness decreases (Mrazek, Smallwood, & Schooler, 2012).

Therefore, when thoughts of past, present, and forthcoming online interactions occur during a task, online vigilance may result in increased mind-wandering. Although to date there is little direct evidence for such vigilance-induced mind-wandering, Stothart, Mitchum, and Yehnert (Stothart et al., 2015) argued that smartphone notifications elicited mind-wandering in participants which was responsible for an impairment in sustained attention. However, they did not measure mind-wandering. In a similar vein, other work suggests mobile technology constantly reminds people of how easily they can communicate with others and that these task-irrelevant thoughts lead to a disruption in task performance (Thornton et al., 2014; Ward et al., 2017). This reasoning can also explain why heavy smartphone users experience higher levels of rumination (Elhai & Contractor, 2018). Last, a moderate amount of mind-wandering episodes have shown to be explicitly about online content and applications (Hollis & Was, 2016). Taken together, we predict that online vigilance is positively related to mind-wandering (H_{1a}).

On the flip side, those high in vigilance should also experience less mindfulness. For instance, individuals with higher levels of excessive social media use are often preoccupied with thoughts about the online world and report lower mindfulness (Sriwilai & Charoensukmongkol, 2016). Similarly, automatic texting behavior, a concept related to the monitoring dimension of online vigilance, negatively predicted facets of mindfulness (Bayer, Dal Cin, Campbell, & Panek, 2016). Thus, it appears those who have a strong mental preoccupation with past, ongoing, or forthcoming online interactions also experience difficulties to focus their attention on the present moment. Consequently, we predict that online vigilance is negatively related to mindfulness (H_{1b}).

Increased mind-wandering and decreased mindfulness present plausible mechanisms that may connect online vigilance to decreased levels of well-being. Whereas mind-wandering has repeatedly shown to be negatively associated with well-being outcomes

(Killingsworth & Gilbert, 2010; Smallwood, O'Connor, Sudbery, & Obonsawin, 2007), mindful individuals, in general, display greater well-being (Frieze & Hofmann, 2016; Gu, Strauss, Bond, & Cavanagh, 2015; Keng, Smoski, & Robins, 2011). Other recent work suggests that a mindful use of instant messaging positively relates to well-being (Bauer, Loy, Masur, & Schneider, 2017). Furthermore, mindfulness mediated the relationship between problematic smartphone use and well-being outcomes (Elhai, Levine, O'Brien, & Armour, 2018). Based on this research and our theoretical assumptions leading to H_{1a} and H_{1b} , we propose that mind-wandering as well as mindfulness act as mediators between online vigilance and psychological well-being.

Thus, we predict direct, negative relationships between mind-wandering and both satisfaction with life (H_{2a}) and affective well-being (H_{2b}), and direct, positive relationships between mindfulness and these well-being indicators (H_{3a} , H_{3b}). Furthermore, we expect an indirect relation between online vigilance and well-being: We predict online vigilance relates negatively to both satisfaction with life and affective well-being via higher mind-wandering (H_{4a} and H_{4b}) and via lower mindfulness (H_{5a} and H_{5b}).

METHOD

Given the recent call to improve the replicability of scientific studies, and in order to reduce false positives (Nosek, Ebersole, DeHaven, & Mellor, 2018), we preregistered the hypotheses outlined above, as well as sample size justification, analyses plan, and exclusion criteria before data collection. Readers can find the preregistration, data, analysis script, and study materials on the Open Science Framework (<https://osf.io/ufyq4/>).

PARTICIPANTS AND PROCEDURE

In total, 497 respondents participated in our online survey hosted by Qualtrics. Participants were students from Radboud University who participated for course credit; in addition, we also employed snowball sampling, that is, posted the survey on Facebook and disseminated it within our personal networks. Participants were invited to participate in a survey about media use and personality. In light of the generally small effect sizes in media effects research (Rains, Levine, & Weber, 2018), we aimed to detect a smallest effect size of interest of $|p| = .15$ (Lakens & Evers, 2014). Thus, for a two-tailed correlation with $\alpha = .05$ to achieve 80% power, we required a sample of 343 participants.

In line with our a priori exclusion criteria, we first removed 112 participants because they did not finish the survey. Second, we followed recommendations on how to obtain high-quality data by excluding participants who did not take the survey seriously, as indicated by an extremely long or short survey time or clicking the same option for each item ("straightlining"). To account for the former, we relied on the Relative Speed Index (RSI), developed by Leiner (2013), which gives an indication of how quickly a participant went through a survey in relation to all other participants. To account for straightlining,

we examined variables with a variance of zero. Accordingly, we excluded 14 participants because they had an RSI > 1.75. Thus, our final sample consisted of 371 participants (70% females) with a mean age of 21.47 ($SD = 5.65$), of whom almost everyone owned a smartphone (369).

MEASURES

ONLINE VIGILANCE

To assess online vigilance, we employed the Online Vigilance Scale, developed and validated by Reinecke et al. (2017). The scale consists of three dimensions (salience, monitoring, reactivity) with four items each. Respondents answered items such as “My thoughts often drift to online content” on Likert-scales, ranging from 1 (*Does not apply at all*) to 5 (*Fully applies*). As suggested by Reinecke et al., the three subscales were aggregated to form an overall indicator of online vigilance. In line with the scale validation of Reinecke et al., the scale displayed high internal consistency ($M = 2.54$, $SD = .72$, $\alpha = .89$).

Mind-Wandering. To measure trait mind-wandering, we employed the commonly used Daydreaming Frequency Scale (Giambra, 1993). The scale consists of twelve items that assess the frequency of absentmindedness in everyday situations and has five different answer options depending on the items, increasing from little to a lot of mind-wandering. For instance, respondents rated items such as “Instead of noticing people and events in the world around me, I will spend approximately...” on a scale from 1 (*0% of my time lost in thought*) to 5 (*50% of my time lost in thought*). The scale displayed excellent internal consistency ($M = 3.08$, $SD = .75$, $\alpha = .92$).

MINDFULNESS

To measure mindfulness, we employed the FFMQ-SF (Bohlmeijer, ten Klooster, Fledderus, Veehof, & Baer, 2011). The scale measures five facets of mindfulness (observe, describe, act aware, nonjudge, nonreact) and consists of 24 items. Respondents rated statement such as “I find it difficult to stay focused on what’s happening in the present moment” on Likert-style ratings ranging from 1 (*never or rarely true*) to 5 (*very often or always true*). The aggregated scale displayed high internal consistency ($M = 3.26$, $SD = .48$, $\alpha = .84$).

SATISFACTION WITH LIFE

We measured the cognitive component of subjective well-being with the Satisfaction with Life Scale (Diener, Emmons, Larsen, & Griffin, 1985). The scale consists of five items, such as “I am satisfied with my life”, that respondents rate on Likert-style scales ranging from 1 (*strongly disagree*) to 7 (*strongly agree*). The scale showed high internal consistency ($M = 4.87$, $SD = 1.17$, $\alpha = .86$).

AFFECTIVE WELL-BEING

We measured the affective component of subjective well-being with the Scale of Positive and Negative Experience (Diener et al., 2010). The scale assesses both positive ($M = 22.69$, $SD = 3.53$, $\alpha = .88$) and negative affect ($M = 15.75$, $SD = 4.18$, $\alpha = .83$) with six items each. Respondents reported how much they experienced affect such as “positive” or “negative” in the past four weeks, and rate those on a Likert-type scale from 1 (*very rarely or never*) to 5 (*very often or always*). By subtracting negative affect from positive affect, we obtained an overall balance measure, with higher scores indicating more positive affect ($M = 6.94$, $SD = 6.99$).

RESULTS

CONFIRMATORY ANALYSES

To test our hypotheses, we estimated a path model with maximum likelihood estimation using the lavaan package (Rosseel, 2012) in R (R Core Team, 2018). We controlled for age and gender in the model. Mardia’s test, Henze-Zirkler’s test and the E-statistic all indicated that our data were not multivariate normal (all $p < .001$). Thus, to deal with the nonnormal distribution, we employed 10,000 bootstrap samples for our models. In addition, all results presented below remained unchanged when we used a robust estimator, namely maximum likelihood estimation with robust standard errors and a Satorra-Bentler scaled test statistic. By employing bootstrapping, we also followed recommendations of Shrout and Bolger (2002) who advise to use bootstrapping to obtain more reliable results for indirect effects. Therefore, we obtained indirect effects by bootstrapping the combined direct effects. To give an example of the combined direct effects, the indirect effect of online vigilance on satisfaction with life via mindfulness was obtained by multiplying the direct effect of online vigilance on mindfulness with the direct effect of mindfulness on satisfaction with life. We determined the criteria for model fit before-hand (for details see our preregistration).

Zero-order correlations are displayed in Table 1. Our original model with all specified paths and no error covariances did not fit the data well, $\chi^2(3) = 48.83$, $p < .001$, CFI = .90, RMSEA = .20, 90%CI[.16, .26], SRMR = .08. Following the steps in our preregistration, we added a covariance between mind-wandering and mindfulness, based on theoretical accounts which predict a moderate relation between the two (Mrazek et al., 2012). This resulted in a good model fit, $\chi^2(2) = 4.01$, $p = .13$, CFI = .99, RMSEA = .05, 90%CI[.00, .13], SRMR = .02. The final model is displayed in Figure 1.

Table 1. Means, standard deviations, and correlations with confidence intervals

Variable	<i>M</i>	<i>SD</i>	1	2	3	4
1. VIG	2.54	0.72				
2. MW	3.08	0.75	.17**			
			[.07, .27]			
3. MF	3.26	0.48	-.31**	-.38**		
			[-.40, -.22]	[-.46, -.29]		
4. LS	4.87	1.17	-.07	-.22**	.47**	
			[-.17, .03]	[-.31, -.12]	[.38, .54]	
5. SPANE	6.94	6.99	-.14**	-.30**	.56**	.66**
			[-.24, -.04]	[-.39, -.20]	[.48, .63]	[.60, .72]

Note. *** indicates $p < .001$; ** indicates $p < .01$. Values in square brackets indicate the 95% confidence interval for each correlation. VIG = Vigilance. MW = Mind-Wandering. MF = Mindfulness. LS = Satisfaction with Life. SPANE = Affective Well-Being.

Supporting H_{1a} , online vigilance was significantly and positively related to mind-wandering ($\beta = .17, p = .002$). Mind-wandering, in turn, did not significantly predict satisfaction with life ($\beta = -.05, p = .336$), rendering no support for H_{2a} , but was significantly and negatively related to affective well-being ($\beta = -.10, p = .040$), supporting H_{2b} .

Supporting H_{1b} , online vigilance was significantly and negatively related to mindfulness ($\beta = -.31, p < .001$). In support of H_{3a} , mindfulness was significantly and positively related to satisfaction with life ($\beta = .46, p < .001$); likewise, supporting H_{3b} , mindfulness was significantly and positively related to affective well-being ($\beta = .53, p < .001$).

Finally, we expected indirect relationships between online vigilance and the two well-being indicators via mind-wandering and mindfulness, respectively. The data did not show a significant indirect relationship of online vigilance and satisfaction with life via mind-wandering (H_{4a} ; $\beta = -.01, p = .354$), nor with affective well-being (H_{4b} ; $\beta = -.02, p = .081$). In contrast, there was a significant negative indirect relationship of online vigilance via mindfulness with life satisfaction (H_{5a} ; $\beta = -.14, p < .001$) and with affective well-being (H_{5b} ; $\beta = -.16, p < .001$).

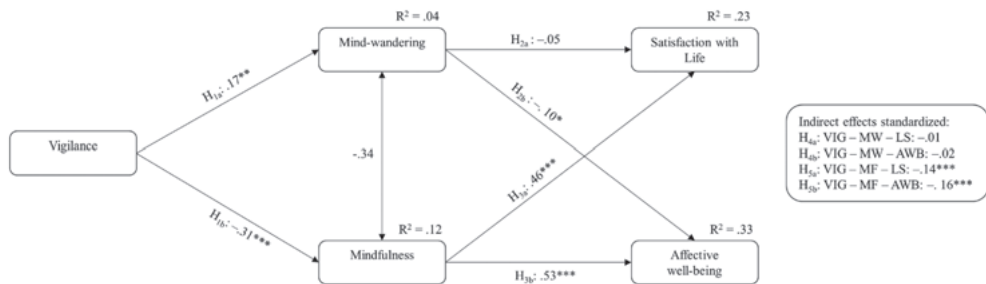


Figure 1. Final path model after adding an error covariance between mind-wandering and mindfulness. All coefficients standardized. * $p < .05$; ** $p < .01$; *** $p < .001$. $\chi^2(2) = 4.01, p = .13$, CFI = .99, RMSEA = .05, 90%CI[.00, .13], SRMR = .02.

EXPLORATORY ANALYSES

To further examine the nature of online vigilance, we explored how each dimension contributed to the proposed mechanism. Consequently, we estimated the same model as above, but split online vigilance into its three dimensions as predictors. The model had mediocre fit, $\chi^2(12) = 37.47, p < .001, CFI = .97, RMSEA = .08, 90\%CI[.05, .10], SRMR = .04$. Interestingly, salience appeared to be the most crucial component of online vigilance, as only salience was directly related to mind-wandering ($\beta = .16, p = .007$) and mindfulness ($\beta = -.24, p < .001$), and indirectly via mindfulness to satisfaction with life ($\beta = -.11, p < .001$) and affective well-being ($\beta = -.13, p < .001$). All other paths were nonsignificant. However, given that these results are exploratory and obtained post-hoc, p-values are not meaningful and the results should be treated as preliminary until independently replicated (De Groot, 2014; Nosek et al., 2018; Wagenmakers et al., 2012). The exploratory model can be found on the OSF.

DISCUSSION

With this study, we address a phenomenon that is becoming increasingly prevalent in society: Technology users report to be in a state of constant alertness, which we refer to as online vigilance. We asked how users experience this online vigilance and investigated whether it relates to decreased well-being due to increased mind-wandering and decreased mindfulness.

Indeed, the results of our study show that those high in online vigilance experienced more mind-wandering and less mindfulness. In the case of mind-wandering, our findings support the notion that a constant preoccupation with online streams of information and communication coincides with more frequent task-unrelated thoughts during the day. Specifically, the correlation is consistent with the assumption of previous experimental work that smartphones may elicit mind-wandering (Stothart et al., 2015; Thornton et al., 2014). In such a view, smartphones represent an omnipresent communication channel to friends and family, which prompts thoughts about what is happening within one's social circle. Therefore, our study shows that being concerned with streams of online information is related to more task-unrelated thoughts, which lends external validity to the assumptions of previous experimental work.

Furthermore, our study demonstrates that those high in vigilance experienced less mindfulness to a considerable degree. Apparently, people constantly devoting attention to past, ongoing, or forthcoming online interactions experience problems to focus on the present moment. Our findings are in line with previous work on phone use and mindfulness. For example, automatic texting behavior does not require an observation of current thoughts and feelings and was thus related to lower mindfulness (Bayer et al., 2016). In particular, automated checking behavior in the form of monitoring and a more or less automatic response to online stimuli in the form of reactivity are in stark contrast

to mindfulness, which is reflected in the moderate to large effect size of the negative correlation between online vigilance and mindfulness.

In addition to finding a link between online vigilance and mind-wandering and mindfulness, our study mostly replicated the links between these constructs and well-being established in previous work. Particularly mindfulness was strongly related to both satisfaction with life and affective well-being, providing further support for the benefits of mindfulness (Frieze & Hofmann, 2016; Gu et al., 2015; Keng et al., 2011). However, contrary to previous findings on the state-level (Killingsworth & Gilbert, 2010; Smallwood et al., 2007), mind-wandering was only weakly related to both well-being outcomes on the trait-level. In addition, given that we investigated this relationship within the path model while controlling for mindfulness, our findings provide support for a recent account claiming that mindfulness mediates the relationship between mind-wandering and well-being (Stawarczyk, Majerus, der Linden, & D'Argembeau, 2012).

More important, the study shows partial support for our prediction that online vigilance relates to well-being through increased mind-wandering and decreased mindfulness. In particular, our results show that mindfulness appears to be the key factor in this mechanism: Online vigilance was indirectly related to both well-being outcomes via lower levels of mindfulness. Although the size of this indirect relationship was rather small, it did explain a significant amount of variance in both well-being outcomes. Moreover, for exploratory reasons we also ran a model that included direct paths from online vigilance to the well-being indicators (available on the OSF). The paths were not significant and model fit was poor. Thus, the total correlation between online vigilance and satisfaction with life was masked by the mediation and not significant in itself. First, this underlines the need to consider mediating mechanisms when investigating the relations between media-related concepts and well-being outcomes. Second, given the direct relation between pathological media use (e.g., problematic Facebook use) and well-being (Marino et al., 2018), our findings are in line with the idea that online vigilance is nonpathological. Mind-wandering, however, did not function as a mediator. Therefore, the role of mind-wandering in the relationship between online vigilance and well-being appears less important as soon as mindfulness is considered simultaneously.

One important limitation of our design is that it only allowed us to investigate these links at a person-level; thus, it did not explicitly test a situational link between smartphone notifications, online vigilance, and the other measures. In addition, some of the effect sizes we observed were relatively small. For instance, online vigilance can only account for a small amount of variance in mind-wandering. This may point to a clear conceptual difference between online vigilance and mind-wandering. More important, there is a need for future work to examine the exact contributions of the specific components of online vigilance to the components of mind-wandering, especially given recent theorizing about the different forms mind-wandering can take (Seli et al., 2018). Likewise, we call for more research on the relation between online vigilance and mindfulness. Our exploratory analyses indicate that salience might be the most potent, and possibly the only predictor of

well-being through decreased mindfulness. However, this fine-grained analysis is post-hoc (Nosek et al., 2018; Wagenmakers et al., 2012), and we call for independent, preregistered replications of this finding.

Taken together, our study examined a potential mechanism of how being constantly vigilant about one's online communication relates to well-being: Those mentally preoccupied with online communication were overall less satisfied with their lives and reported less affective well-being when they also experienced reduced mindfulness. However, this mechanism does not mean that online vigilance has negative consequences per se. On the contrary, online vigilance has the potential to increase well-being by making access to social support, enjoyable content, and social gratifications cognitively salient and available (Karapanos et al., 2016; van Koningsbruggen et al., 2017). Our results imply that potential positive effects on well-being may be contingent on whether it reduces mindfulness. While our findings should be interpreted as preliminary due to the cross-sectional nature of our design, they give a first indication of the importance to advance research on the topic of online vigilance.

The Relationship Between Online Vigilance and Affective Well-Being in Everyday Life: Combining Smartphone Logging with Experience Sampling

This chapter is currently under review as Johannes, N., Meier, A., Reinecke, L., Ehlert, S., Setiawan D. N., Walasek, N., Dienlin, T., Buijzen, M., & Veling, H. (under review). The Relationship Between Online Vigilance and Affective Well-Being in Everyday Life: Combining Smartphone Logging with Experience Sampling.



ABSTRACT

Through communication technology, users find themselves constantly connected to others to such an extent that they routinely develop a mindset of connectedness. This mindset has been defined as *online vigilance*. Although there is a large body of research on media use and well-being, the question of how online vigilance impacts well-being remains unanswered. In this preregistered study, we combine experience sampling and smartphone logging to address the relation of online vigilance and affective well-being in everyday life. Seventy-five Android users answered eight daily surveys over five days ($N = 1615$) whilst having their smartphone use logged. Thinking about smartphone-mediated social interactions (i.e., the salience dimension of online vigilance) was negatively related to affective well-being. However, it was far more important whether those thoughts were positive or negative. No other dimension of online vigilance was robustly related to affective well-being. Taken together, our results suggest that online vigilance does not pose a serious threat to affective well-being in everyday life.

Research on the question of how communication technology affects the well-being of users has accumulated rapidly in recent years (Meier, Domahidi, & Günter, in press). Results so far illustrate that the use of such technology has small (Heffer, Good, Daly, MacDonell, & Willoughby, 2019; Orben, Dienlin, & Przybylski, 2019; Orben & Przybylski, 2019b, 2019a) and possibly nonlinear effects on well-being (Przybylski & Weinstein, 2017).

These findings all refer to communication technology use at the *behavioral level* (e.g., by investigating “screen time”). However, such approaches neglect that people are constantly connected to others *psychologically* through their smartphones. This connection has led smartphone users to develop a mindset of constant connectivity, a phenomenon recently defined as online vigilance (Klimmt et al., 2018; Reinecke et al., 2018). Importantly, in contrast to other theoretical approaches such as problematic internet use (Kardefelt-Winther et al., 2017), online vigilance refers to a non-pathological form of constant psychological connectedness to online content and communication. People high in online vigilance are perpetually aware of streams of mediated communication in daily life. This awareness has the potential to contribute to users’ well-being, for example, when it takes the form of perceived social support (Domahidi, 2018; Reinecke, 2018). Yet many users experience such a mindset as bothersome and conflicting with personal goals and obligations (Mihailidis, 2014; Näsi & Koivusilta, 2013). Thus, the constant cognitive preoccupation resulting from online vigilance may impair individuals’ psychological well-being (Reinecke, 2018). Therefore, in addition to focusing on *usage of technology*, there is a need for research investigating how a *mindset of connectedness due to technology* relates to well-being.

Despite the proliferation of smartphone use and resulting opportunities to develop online vigilance, such research examining the relation between online vigilance and well-being is scarce. Moreover, previous work assessed the relationship between online vigilance and well-being on the basis of cross-sectional self-reports (Johannes et al., 2018). However, cross-sectional approaches cannot analyze situational within-person processes (Hamaker, Kuiper, & Grasman, 2015), while self-reports provide only unreliable accounts of actual behavior (e.g., Scharrow, 2016). As a result, the current literature can only provide a coarse picture, leaving room for many confounding factors and limiting insights into the mechanisms of how online vigilance relates to well-being.

To address these shortcomings, there is a need for research that investigates this relation in the situational contexts of everyday life. Thus, the present study extends prior work by combining experience sampling with behavioral data (i.e., objective smartphone logging). This approach makes several contributions. First, it addresses methodological limitations of prior work. Second, it allows us to get a better understanding of the online vigilance construct on the state level. Third, instead of expecting online vigilance to display a uniform relation with well-being, we relied on a more fine-grained theoretical approach and investigated how individual dimensions of online vigilance relate to well-being. Fourth, rather than focusing on online vigilance concerning general online media use, we focus on users’ mindset towards *smartphone-mediated social interactions*, because social interaction

is a core function of smartphone use (e.g., Klimmt et al., 2018). Together, the current study advances our understanding of how constant cognitive connectedness in the form of online vigilance relates to well-being.

ONLINE VIGILANCE AND WELL-BEING

Via their smartphones, users are now permanently connected to social contacts (e.g., Mihailidis, 2014). This *technological* connectedness can lead users to internalize a *psychological* connectedness, such that users are constantly aware of ongoing streams of mediated communication and interaction. Reinecke, Klimmt, et al. (2018) have introduced the concept of online vigilance to describe individual differences in this connectedness mindset. Online vigilance is reflected in three features of users' psychology: "(1) their *cognitive orientation* to permanent, ubiquitous online connectedness; (2) their *chronic attention* to and continuous integration of online-related cues and stimuli into their thinking and feeling; and (3) their *motivational disposition* to prioritize options for online communication over other (offline) behavior" (Reinecke et al., 2018, p. 2). These features are expressed in three dimensions of online vigilance: salience, reactivity, and monitoring. *Salience* refers to the frequency and intensity of thoughts about online streams of communication and interaction, thus representing the cognitive component of online vigilance. *Reactivity* refers to the motivational component of online vigilance, specifically the sensitivity to smartphone cues and how responsive the user is to them, even when this requires postponing ongoing offline activities. *Monitoring* describes the attentional component of online vigilance, specifically to what extent users observe their online sphere, expressed in how often a user checks their mobile device proactively without being prompted by a notification.

The dimensions of online vigilance can exert different influences on well-being, depending on whether the dimensions manifest themselves in thoughts and behavior that are conducive to the current task or not (Reinecke, 2018). Supporting such a view, the literature on media use and well-being suggests that goal-directed, purposeful social interactions via media can enhance social gratifications (Bayer et al., 2018; Burke & Kraut, 2016; Jung & Sundar, 2018) and contribute to well-being (Domahidi, 2018; Trepte, Dienlin, & Reinecke, 2015). In contrast, passive use or technology use as procrastination can have negative consequences for well-being (Meier, Reinecke, & Meltzer, 2016; Reinecke & Hofmann, 2016; Verduyn et al., 2015). Analogously, it is not pertinent to assume that the dimensions of online vigilance generally have a positive or negative relation to well-being. Just as mediated communication that conflicts with other goals presents a challenge for well-being, we expect different mechanisms to connect each dimension of online vigilance to well-being. Specifically, we suggest dimensions of online vigilance to relate negatively to well-being if higher levels of the specific dimension (salience, reactivity, monitoring) are expressed in thoughts and behaviors that represent an *interference* with people's lives or higher-order goals.

If the online vigilance dimensions take the form of interferences in everyday life, they should exert their influence on a situational level. In other words, it is unlikely that such interferences instantly influence how people generally evaluate their lives. Satisfaction with life, for instance, has shown to be rather stable and only displays gradual changes (Diener et al., 2018). Instead, if dimensions of online vigilance take the form of situational thoughts and behaviors, they should relate to situational *affect*. Therefore, we expect the online vigilance dimensions to relate to affect as a transient well-being construct that is sensitive to small, moment-to-moment changes in individuals' well-being (Wilhelm & Schoebi, 2007). Such affective well-being has been shown to be affected by intraindividual variations in social media behavior (Bayer et al., 2018).

SALIENCE

So far, there is no evidence linking situational online vigilance to affective well-being. However, on the trait level, there is initial evidence that online vigilance indeed can take the form of interference. For example, online vigilance related negatively to well-being through decreased mindfulness. This indirect association appears to be mostly driven by the *saliency* dimension (Johannes et al., 2018). In other words, thoughts about mediated interactions were most detrimental when they distracted from the current moment. Such a mechanism can also explain why online vigilance has been linked to perceived stress (Reinecke et al., 2018): Saliency in the form of interfering thoughts could be perceived as stressful. The prominent role of the saliency dimension is not surprising, given that absentmindedness in the form of mind-wandering has been shown to be negatively related to well-being (e.g., Smallwood et al., 2007). Such task-irrelevant thoughts may distract from the current moment, thereby decreasing well-being (Franklin et al., 2013; Killingsworth & Gilbert, 2010), particularly when pondering about the past (Spronken, Holland, Figner, & Dijksterhuis, 2016). Saliency and mind-wandering are conceptually related. Whereas mind-wandering manifests in conscious thoughts, saliency encompasses both conscious thoughts and unconscious preoccupation with online communication. When saliency takes the form of conscious, task-irrelevant thoughts, we expected such thoughts to interfere with experiencing the current moment. In line with such a prediction, when people use their smartphone in a mindful rather than in a mindless way, these negative effects can reverse (Bauer et al., 2017). Taken together, we expected thoughts about mediated interactions (i.e., situational saliency) to distract from the immediate environment. Hence, we predicted that saliency is related negatively to affective well-being (H_1).

However, such a view assumes that all experiences of saliency represent task-irrelevant thoughts. Yet thoughts can be of different valence. Thinking of desired experiences can alleviate boredom (Eastwood, Frischen, Fenske, & Smilek, 2012) and induce goal-setting for the distant future (Mooneyham & Schooler, 2013), which is experienced as positive (Spronken et al., 2016). Likewise, social daydreaming, over time, can lead to positive affect and less loneliness (Poerio et al., 2016). Thus, the valence of thoughts plays a role in the effect of mind-wandering on well-being, with interesting and positive thoughts relating

positively, but other thoughts relating negatively to well-being (Franklin et al., 2013). Such a mechanism likely applies to salience as well, as the online vigilance construct allows for both negative and positive effects on well-being (Reinecke, 2018). Thinking about one's social network and the support it provides, analogous to how interactions with close ties via technology can increase well-being (Burke & Kraut, 2016), may thus alleviate the negative, distracting effect of salience. In other words, thoughts about online streams of communication and interaction may distract from the current environment. If those thoughts are positive, however, they might compensate for the distraction. Thus, we predicted that the valence of thoughts moderates the relationship between salience and affective well-being, such that the negative relationship is stronger with negative valence, but weaker with positive valence (H_4)¹.

Furthermore, not only is there no work documenting the frequency of thoughts about mediated interactions; it is also unclear how frequently these thoughts occur compared to thoughts about face-to-face interactions. Similarly, a comparison of the valence of thoughts about mediated interactions versus thoughts about face-to-face interactions would provide valuable insights into the nature of the salience dimension. That is, without such descriptive information, it is difficult to assess whether salience differs from any other sort of mental preoccupation with interpersonal communication. For descriptive information, and to provide an exploratory comparison, we thus investigate the frequency and valence of both thoughts about online as well as face-to-face interactions.

REACTIBILITY

In the case of the reactivity dimension, previous research shows that people appear to be extremely sensitive to smartphone notifications. Many users respond to notifications almost instantly; even if their phone is in silent mode they check notifications within minutes (Pielot, Church, et al., 2014). Such responsiveness to smartphone cues may come at the cost of increased stress. Smartphone users high in reactivity who attend instantly to notifications will routinely interrupt other tasks (Mehrotra, Pejovic, Vermeulen, Hendley, & Musolesi, 2016). Smartphone notifications can serve as connection cues that automatically capture attention for users with high reactivity (Bayer et al., 2015). This can interrupt current tasks (Stothart et al., 2015). Crucially, smartphone-induced interruptions can ultimately lead to high communication load and stress (e.g., Reinecke, Aufenanger, et al., 2017). For instance, increasing smartphone cues led to feelings of social pressure (Halfmann & Rieger, 2019). Conversely, minimizing notification alerts has shown to decrease inattention, which was responsible for increased well-being (Kushlev et al., 2016). We thereby reasoned that people high in reactivity would be more responsive to notifications, resulting in lower well-being. Hence, we predicted reactivity relates negatively to affective well-being (H_2).

¹ We label hypotheses from H_1 to H_4 to stay consistent with the labeling used in our preregistration; see Method.

MONITORING

Last, people high in monitoring display more quick checks of their smartphone, unprompted by a notification. Such checks have shown to occur frequently and often manifest themselves in the form of habits (Hintze, Hintze, Findling, & Mayrhofer, 2017; Oulasvirta, Rattenbury, Ma, & Raita, 2012). Whereas monitoring has the potential to remind people of their social network (Domahidi, 2018), unprompted checks often do not serve an explicit goal; instead, monitoring regularly takes the form of non-purposeful checks (Oulasvirta et al., 2012). Monitoring can then be understood as a specific form of mindless media use, which has shown to relate negatively to well-being (Meier et al., 2016; Verduyn et al., 2015). Repeatedly checking in to online streams of communication and interaction without an explicit communication goal represents a distraction from the current moment. Evidence for such a prediction comes from research showing that checking email less frequently was associated with increased well-being (Kushlev & Dunn, 2015). The distracting mechanism behind the hypothesized effect of monitoring is distinct, though, from that of salience or reactivity. Both salience and monitoring are self-initiated and not necessarily prompted by notifications, yet salience refers to the cognitive components of online vigilance, whereas monitoring is expressed behaviorally. Furthermore, both reactivity and monitoring are expressed behaviorally, but reactivity is exclusively prompted externally, whereas monitoring can be prompted both internally and externally (Klimmt et al., 2018; Reinecke et al., 2018). Thus, monitoring should result in distraction from the current moment, independent of the distraction caused by salience or monitoring. Therefore, we predicted monitoring relates negatively to affective well-being (H_3).

THE CURRENT STUDY

With the current study, we had two central goals. Theoretically, we investigated the relationship between online vigilance and affective well-being in a new context, namely, within smartphone-mediated social interactions on the state level. In our study, we focused on thoughts about social interactions as well as actual interactions via the smartphone. We chose this focus because the technological connectedness smartphones afford is mostly due to their social features. For example, instant messaging apps or social media apps afford users constant connection to others. In other words, this *technological* connectedness to social contacts is the primary source of *psychological* connectedness (Klimmt et al., 2018; Reinecke et al., 2018).

Methodologically, we aimed at addressing several shortcomings in the literature. Traditionally, the majority of research on smartphone use and well-being has relied on self-reports of usage behavior (for a recent critique, see Ellis, 2019). However, there is increasing evidence that people are poor estimators of their phone use, casting doubt on the validity of self-reported screen time (Ellis et al., 2019; Scharkow, 2016; Wilcockson, Ellis, & Shaw, 2018). Moreover, it is questionable whether general trait measures are predictive

of behavior in the moment (e.g., Masur, 2018). To address these limitations, in the present study we combined smartphone logs as behavioral indicators of online vigilance with experience sampling of situational self-report data of the three online vigilance dimensions and of affective well-being. To our knowledge, there is only one study that combined objective smartphone use with experience sampling. Katevas, Arapakis, and Pielot (2018) found that phone use at night, not general phone use, negatively predicted well-being. However, the analysis aggregated phone use and well-being per day, rather than predicting each instance of reported well-being with preceding phone use variables. For one, this demonstrates the need for more fine-grained analyses. Second, their analysis was data driven, demonstrating the need for confirmatory, hypothesis testing work.

The need for confirmatory research is particularly important given that many findings in the Social Sciences do not replicate (e.g., Camerer et al., 2018), likely due to undisclosed flexibility in data analysis (e.g., Nelson & Simmons, 2018). These criticisms have led to calls for confirmatory research that explicates all hypotheses and analysis steps before the data are collected in so-called *preregistrations* (Nosek et al., 2018). As a consequence, the effect of media use on well-being might have been overestimated so far (e.g., Orben & Przybylski, 2019b). We thus preregistered this project, thereby restricting flexibility in confirmatory data analysis in an attempt to increase the reliability of our findings.

METHOD

We preregistered all hypotheses, operationalizations, exclusion criteria, and analysis steps on the Open Science Framework (OSF, https://osf.io/xa74g/?view_only=fc8243cd484c4426923f67055010c286), where readers can also find all materials, data, and analysis scripts (https://osf.io/n6d8k/?view_only=443d7d4c7c174b4bb128a221896c8945).

PARTICIPANTS

Due to our analytical approach and a lack of previous research, power calculations were difficult. We followed the pragmatic recommendation to recruit as many participants as we had resources for (Albers & Lakens, 2018). Thus, we preregistered to collect data from 200 participants or to end collection at a preregistered date (i.e., September 1, 2018). Between February and September 2018, we recruited 111 students from a Dutch university. Participants had to be undergraduate students, proficient in English, between 18 and 30 years old, use an Android phone, and had to use at least one of the following social media apps daily: WhatsApp, Facebook, Facebook Messenger, Instagram, Snapchat, Twitter, SMS/Messenger. Participants could choose between receiving money or credits as reimbursement for their participation. We only recruited Android users because smartphone logging only worked on that operating system. Both forms of reimbursement followed an incentive scheme, such that participants received 5€ for the intake session

and an additional 0.50€ for every two surveys they filled out. In total, participants could earn up to 15€ or an equivalent amount of course credit.

We had to exclude a substantial number of participants because of technological issues. Because of these issues, the logging app could neither log phone use nor send surveys to several devices. For a detailed explanation and all decisions made during data collection, see the Online Supplementary Materials (OSM) on the OSF project. We retained the data from 77 participants who had complete logging and survey data. According to our preregistered exclusion criteria, we excluded one additional participant for answering less than eight surveys (i.e., 20% of 40 scheduled surveys). Finally, one participant received 69 non-duplicate surveys due to the technological issues described above. Because we could not be sure whether we could trust the data from this participant, we excluded the data.

In total, we retained data from 75 participants (53 female, 21 male, 1 preferred not to indicate gender) with a typical age range for undergraduate students ($M_{age} = 21.89$, $SD_{age} = 2.48$). All participants indicated that they used at least one of the social apps of interest daily. All participants gave informed consent; the study had approval from the ethics board (ECSW-2C17-059R1).

PROCEDURE

Participants arrived in the lab to participate in a study titled “Smartphone use in everyday life”. They were informed that the goal of the study was to assess how people use their smartphones in daily life. They received further information about the nature of the study, the logging procedure, how many surveys they would receive daily, and the survey questions. Next, participants installed the logging app with the help of the researcher. We used the app PACO (www.pacoapp.com) for data collection. The app can log phone use and distribute experience sampling surveys, is open source, and free to use, but works only for Android devices, as iOS does not allow phone logging. The app recorded when the screen was unlocked or locked, when and what app was used, and when and from what app notifications arrived.

After the installation, participants took an intake survey that assessed trait measures and demographic information. Afterwards, participants were free to choose a time frame of five consecutive working days. Such a time frame has shown to be representative of people’s typical phone use (Wilcockson et al., 2018). On these days, they received eight daily surveys between 09:00 and 21:00. We chose working days because affective well-being is generally different between weekdays and weekends (Helliwell & Wang, 2014). Surveys were sent at semi-random intervals, with at least 45 minutes between surveys. Due to the time frame, some participants received less than the total 40 surveys ($M = 36.33$, $SD = 3.81$), as some of them turned on their phones only later in the day. Surveys contained 18 short questions and took less than a minute; participants had a five-minute time window to respond to surveys. Response rate to surveys was 59.6% ($SD = 17.5\%$). We also collected experiences about the study in an exit survey, the results of which can be found in the OSM on the OSF.

MEASURES

STATE SALIENCE

We chose self-reports instead of a behavioral measure for salience because self-reports are the most adequate measure of thought frequency (Mooneyham & Schooler, 2013). To assess salience on the state level, participants answered one item (“In the last half an hour, how much were you thinking about mediated interactions (e.g., phone calls, WhatsApp messages, Facebook likes, Instagram posts etc.)?”) on a scale from 1 (*not at all*) to 7 (*a lot*), $M_{raw} = 3.56$, $SD_{raw} = 1.93$. During the set-up process, participants were provided with a short explanation of mediated interactions, stating that the term did not refer to face-to-face interactions but to any form of contact participants had with others via their smartphone, tablet, laptop, etc. We informed participants that mediated interactions did not refer only to talking or texting with others, but also more general forms of contact such as receiving or providing a like or a comment. For exploratory purposes, participants also answered the same question about face-to-face interactions (“In the last half an hour, how much were you thinking about face-to-face interactions?”, $M_{raw} = 3.84$, $SD_{raw} = 2.02$). Valence of situational thoughts about social interaction. Participants indicated the pleasantness of thoughts about mediated interactions ($M_{raw} = 4.84$, $SD_{raw} = 1.21$) and face-to-face interactions ($M_{raw} = 5.11$, $SD_{raw} = 1.23$) with one item, respectively (“How pleasant were those thoughts about mediated/face-to-face interactions?”) on a scale from 1 (*unpleasant*) to 7 (*pleasant*).

STATE REACTIBILITY

Originally, we operationalized reactivity as the average time participants took to open a notification from a social app (i.e., WhatsApp, Facebook, Facebook Messenger, Instagram, Snapchat, Twitter, SMS/Messenger). However, due to the technical issues described above, only 37.8% of days had at least one notification logged, which would have resulted in an inordinate amount of missing data for the reactivity measure. To account for possible problems with logging, we preregistered a decision tree, specifying how we would proceed in case of missing data or other problems. Thus, we followed our a priori decision tree and operationalized reactivity as the time between receiving a survey and opening the survey ($M_{seconds} = 64.50$, $\bar{x} = 26.00$, $SD = 71.23$). We also assessed self-reported reactivity with one item (“In the last half an hour, when I received an online message, I immediately gave it my full attention.”) developed by Reinecke, Klimmt et al. (2018). Participants indicated their agreement on a scale from 1 (*strongly disagree*) to 7 (*strongly agree*), ($M_{raw} = 3.72$, $SD_{raw} = 1.98$).

STATE MONITORING

Originally, we operationalized monitoring as the amount of checks before the survey was opened, with a check defined as any sequence of unlocking and locking the screen again without having been preceded by a notification from a social app. However, such unlocks barely occurred, most likely because the app did not log if the screen was turned on without

being unlocked. Thus, we followed our preregistered decision tree and operationalized monitoring as the total time social apps were used in the 30 minutes before the survey was opened ($M_{seconds} = 122.43$, $\bar{x} = 47.00$, $SD = 190.33$). We also assessed self-reported monitoring with one item (“In the last half an hour, I was constantly monitoring what was happening online.”) developed by Reinecke, Klimmt et al. (2018). Participants indicated their agreement on a scale from 1 (*strongly disagree*) to 7 (*strongly agree*), ($M_{raw} = 2.92$, $SD_{raw} = 1.83$).

STATE AFFECTIVE WELL-BEING

Participants indicated how they currently felt on a mood scale (Wilhelm & Schoebi, 2007). The items were preceded with “At this moment, I feel”, followed by six mood dichotomies (e.g., “tired-awake”) from the original scale plus an additional dichotomy (“depressed-happy”) to explicitly capture happiness. Items were aggregated per survey to form a mean index ($M_{overall} = 5.02$, $SD_{overall} = 1.03$).

TRAIT MEASURES

We measured four traits in the intake survey for exploratory analysis. The full details of their measurement can be found in the OSM. To measure online vigilance as an individual difference variable, participants filled out the Online Vigilance Scale (Reinecke et al., 2018), $M = 2.85$, $SD = 0.64$, $\alpha = .86$. To assess the smartphone checking habit of participants, we employed an established 12-item measure of habit strength in the intake survey (Verplanken & Orbell, 2003) and added an item reflecting the lack of intentionality, following suggestions of previous work on smartphone habits (Bayer & Campbell, 2012), $M = 4.85$, $SD = 0.82$, $\alpha = .86$. We assessed the affective component of well-being with the scale of positive and negative experience developed by Diener et al. (2010), $M = 8.73$, $SD = 5.62$. We assessed the evaluative component of well-being with the satisfaction with life scale (Diener et al., 1985), $M = 5.02$, $SD = 0.98$, $\alpha = .80$.

RESULTS

MAIN EFFECTS MODEL

Following our preregistration, we computed two models. The main effects model, testing the situational relationships specified in $H_1 - H_3$, predicted affective well-being from self-reported salience, behavioral reactivity, and behavioral monitoring and was run on the full data. The moderator model, testing H_4 , predicted well-being from the same variables as the main effects model, but also included valence of thoughts related to mediated social interactions as well as an interaction term between salience and valence of thoughts. The interaction model was run on a subset of the data, because the moderator, thought valence, did occur only when people indicated non-zero levels of salience. Because both

models were based on the same data, we adjusted for multiple testing by applying a Bonferroni correction, such that we regarded parameters to be significant if they were below $\alpha = .025$. There was substantial variation of affective well-being between and within participants, $ICC = .38$, demonstrating that multilevel modelling was appropriate. We conducted all analyses in *R* (version 3.5.2, R Core Team, 2018); data wrangling and visualization were done with *tidyverse* (version 1.2.1, Wickham, 2017). To account for the nested nature of our data, we ran mixed-effects models with the *lme4* package (version 1.1-21, Bates, Maechler, Bolker, & Walker, 2015). We followed recommendations to control Type I errors by employing a maximal random effects structure (Barr, Levy, Scheepers, & Tily, 2013). That is, in the original preregistered models we specified random intercepts and random slopes for each predictor for participants nested in days. In addition, we preregistered to person mean-center all continuous predictors. To obtain p -values, we computed bootstrapped Likelihood Ratio Tests using the *mixed* function (*afex* package, version 0.20-2; Singmann, Bolker, Westfall, & Aust, 2018).

Table 1. Parameters and estimates of the two confirmatory mixed-effects models

Predictors	<i>B</i>	<i>SE</i>	<i>t</i>	β	<i>p</i>	R^2_F	R^2_R
Main effects model						.004	.38
Saliency	-0.004	0.03	-0.14	-.01	.97		
Reactivity	0.023	0.02	0.95	.03	.48		
Monitoring	-0.06	0.02	-2.67	-.06	.008		
Interaction model						.05	.44
Saliency	-0.089	0.03	-3.00	-.11	.006		
Valence	0.169	0.03	5.13	.20	.001		
Reactivity	0.034	0.02	1.47	.04	.171		
Monitoring	-0.052	0.02	-2.43	-.06	.042		
Saliency x Valence	0.054	0.02	2.23	.06	.086		

Note. Beta coefficients are based on a model with standardized predictors and outcomes. The other estimates are based on the preregistered models with standardized predictors but non-standardized outcomes. Hence, beta coefficients should be seen as an indicator of the effect size, not a direct translation. R^2_F denotes the variance explained by fixed factors. R^2_R denotes the variance explained by both fixed and random factors.

The model did not immediately converge. We followed current best practices and removed day as a grouping factor. The model still yielded a warning, but at an acceptable tolerance level. For full details on the modelling procedure, see the OSM on the OSF. Model diagnostics, such as residuals and checks for heteroscedasticity, displayed acceptable fit. Of the three predictors, only monitoring negatively predicted well-being, $PBtest = 7.02$, $p = .008$.

To obtain an indication of the standardized effect size, we also standardized the outcome variable (not preregistered). The standardized effect size was small ($\beta = -.06$); see Table 1 for all parameter estimates. We called the *r.squaredGLMM* function to obtain *Pseudo R*² for mixed-models (*MuMIn* package, version 1.40.4, Barton, 2018), displaying $R^2 = .004$ for the variance explained by fixed factors, and $R^2 = .38$ for the variance explained by both fixed and random factors. We conducted several robustness checks (not preregistered, see OSM). Excluding an outlier yielded a nonsignificant (at $\alpha = .025$) estimate of the effect of monitoring, $PBtest = 6.63$, $p = .035$. Because of the convergence issues, we also estimated a Bayesian mixed-effects model with the *brms* package (version 2.2.0, Bürkner, 2017), which estimated parameters very close to those of the frequentist model, increasing our trust in the estimates. The model also supported our suspicion that the monitoring effect is not robust.

INTERACTION MODEL

Confirming that our model was too complex, the original model with an added interaction term yielded a convergence error, as the number of observations was smaller than the number of random effects. We followed the same steps as with the main effects model. In addition to removing the weekday grouping, we also had to remove the correlation between random effects (Barr et al., 2013); the resulting singularity warning was within an acceptable level of tolerance. Model diagnostics displayed good fit.

Supporting our conclusion that the effect of monitoring was not robust, the effect was not significant (at $\alpha = .025$) and small, $PBtest = 5.89$, $\beta = -.056$, $p = .042$, when also accounting for the valence of salience and its interaction with salience. In this model, salience was significantly and negatively related to well-being at average valence, $PBtest = 8.62$, $\beta = -.11$, $p = .006$. However, valence was positively related to well-being at average salience with a larger effect size, $PBtest = 22.83$, $\beta = .21$, $p = .001$. Although in the expected direction, their interaction failed to reach significance, $PBtest = 4.85$, $\beta = .06$, $p = .086$. Thus, salience and the valence of salience were independently related to well-being. Those effects were evident in a larger $R^2 = .05$ for the variance explained by fixed factors, and $R^2 = .44$ for the variance explained by both fixed and random factors. Exploratory robustness checks and Bayesian models demonstrated that the interaction model was robust and displayed better fit than the main effects model.

PREREGISTERED EXPLORATORY ANALYSES

Following our preregistration, we also explored correlations between the person-aggregated state-level measures and trait measures. As those correlations were entirely exploratory, p -values are not meaningful (Gelman & Loken, 2013), which is why we only present and interpret effect sizes, see Figure 1. Interestingly, the correlations between the behavioral indicators and the online vigilance trait were relatively low ($r < .21$), compared to moderate correlations between online vigilance and the self-reported state-level vigilance

dimensions ($r > .34$). Trait online vigilance was only weakly correlated to the trait well-being measures ($r > -.10$), but was strongly related to smartphone habits ($r = .64$).

On the state-level, the self-reported items correlated highly with each other ($r > .68$). On the aggregate level, behavioral monitoring was moderately related to both self-reported reactivity and self-reported monitoring ($r > .37$), whereas behavioral reactivity was not correlated to any other variable, except trait online vigilance.

However, these relationships must be interpreted with caution, as aggregating the state-level variables can only provide coarse estimates. To explore the relationship between self-reported and behavioral measures of salience and monitoring in a more fine-grained manner, we ran two maximal mixed-effects models (without the weekday grouping). In line with the overall correlation, self-reported reactivity was a poor predictor of behavioral reactivity, $t(1) = -5.53, SE = 3.18, \beta = -.14, 95\%CI = [-.19, -.09]$. That means that a higher self-reported attention to notifications was weakly related to responding faster to surveys. Self-reported monitoring was a stronger predictor of behavioral monitoring, $t(1) = 8.27, SE = 6.17, \beta = .31, 95\%CI = [.26, .37]$.

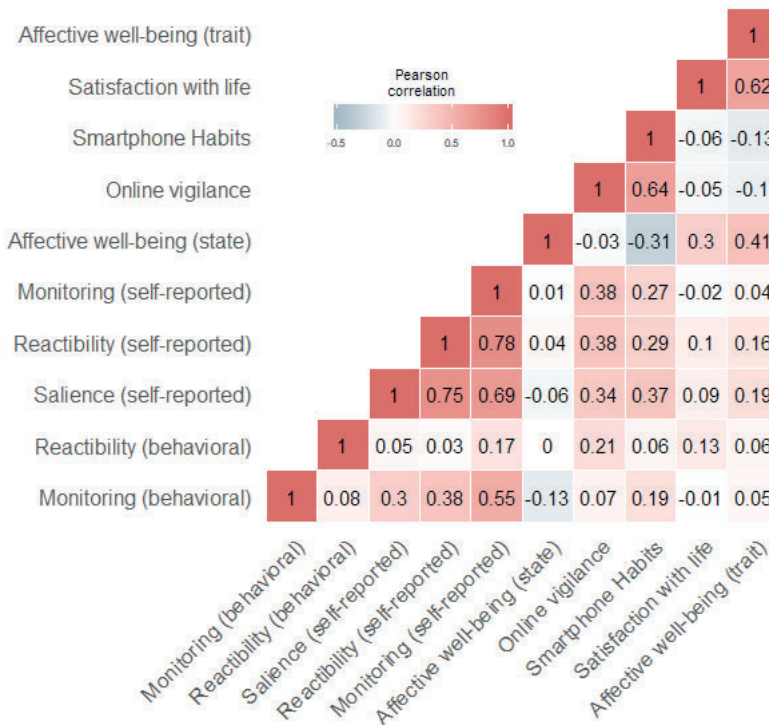


Figure 1. Heat map of correlations between person-aggregated state-level measures and trait measures. The numbers in each field represent Pearson’s correlation coefficient. The strength and direction are visualized with colors.

FURTHER EXPLORATORY ANALYSES

We were also interested in whether self-reported indicators of reactivity and monitoring would predict well-being better than behavioral indicators. To that end, we predicted standardized well-being with self-reported standardized salience, reactivity, and monitoring in a maximal mixed-effects model. All estimates were extremely close to zero (all $\beta < |.04|$), yielding no evidence of an effect of any predictor. See the OSF page for the full details of this analysis.

Inspecting the data showed that there were differences in the frequency and valence between thoughts about mediated interactions and face-to-face interactions. Thoughts about mediated interactions occurred less frequently ($M = 3.52$, $SD = 0.96$) than thoughts about face-to-face interactions ($M = 3.83$, $SD = 1.06$), $t(74) = -2.74$, $d_z = -0.32$. Similarly, valence of thoughts about mediated interactions was more negative ($M = 4.21$, $SD = 0.80$) than that of thoughts about face-to-face interactions ($M = 5.08$, $SD = 0.72$), $t(74) = -7.49$, $d_z = -0.86$.

Consequently, we were interested in whether thoughts about and valence of face-to-face interactions were a stronger predictor of well-being than thoughts about and valence of mediated interactions. A maximal mixed-effects model (no day grouping, no random correlations) showed an interesting pattern of results. Corroborating our previous analyses, salience frequency was negatively related to affective well-being, $t(76.82) = -4.18$, $SE = 0.04$, $\beta = -.15$, $95\%CI = [-.21, -.08]$, whereas the valence of those thoughts displayed a positive relation, $t(72.28) = -5.84$, $SE = 0.03$, $\beta = .19$, $95\%CI = [.13, .26]$. Surprisingly, the frequency of thoughts about face-to-face interactions was not meaningfully related to affective well-being, $t(77.28) = 1.03$, $SE = 0.03$, $\beta = .03$, $95\%CI = [-.04, .10]$; however, their valence exhibited a positive relation to affective well-being, $t(1041.43) = 6.14$, $SE = 0.03$, $\beta = .20$, $95\%CI = [.13, .27]$. Hence, it appears that the frequency of thoughts about mediated interaction is problematic compared to the frequency of thoughts about face-to-face interactions. However, the relation of the valence of both types of thoughts with affective well-being appears more important and of equal size for mediated and non-mediated interactions.

DISCUSSION

With smartphones becoming a central part of people's lives, users report to be in a state of constant alertness (Mihailidis, 2014). This alertness has been defined as online vigilance (Klimmt et al., 2018; Reinecke et al., 2018). To date, it is largely unclear whether this constant orientation toward mediated communication has consequences for well-being. In this study, we address this question. Specifically, we asked whether vigilance is related to well-being in people's everyday lives. To that end, we combined behavioral logging data with momentary self-reports, going beyond the deterministic behavioral view of connectedness as merely expressed in screen time (e.g., Orben & Przybylski, 2019a). Instead, our study focused on the psychological internalization of connectedness,

measured with rarely used real-time logging (Ellis, 2019). Consequently, we provide a thorough test of so far unstudied episodic fluctuations in online vigilance and their relation to situational affective well-being.

Overall, our results suggest small to negligible situational relations between the three dimensions of online vigilance and well-being. These small effect sizes are in line with recent work on the relation between screen time and well-being (Heffer et al., 2019; Orben et al., 2019; Orben & Przybylski, 2019b, 2019a; Przybylski & Weinstein, 2017) and with small effect sizes in the field in general (Rains et al., 2018). Valence of thoughts about mediated communication was by far the strongest predictor of well-being. The more positively participants thought about mediated interactions in the past half an hour, the better they felt in the current moment. This effect was rather large in comparison to the typically small effect sizes in media effects research (Rains et al., 2018). This finding presents a preregistered replication of previous research showing that thoughts about one's social network and positive, interesting thoughts in general can increase positive affect (Franklin et al., 2013; Poerio et al., 2016). More importantly, we extend previous research by providing evidence that the effect of valence is not limited to thoughts about offline interactions, but also applies to thoughts about mediated interactions. This corroborates accounts arguing for smartphones as an ever-present reminder of one's social network (Carolus et al., 2018).

The role of valence becomes particularly important in light of the negative relation of salience to well-being. In line with our predictions, more thoughts about mediated interactions were associated with slightly decreased well-being. This emphasizes the importance of the salience dimension when investigating the relation of the online vigilance construct with well-being. Similar to previous work on the trait level, where salience was the most important predictor (Johannes et al., 2018), salience on the state level was the only dimension of online vigilance robustly predicting well-being. This is in line with previous work showing mind-wandering on the state level to be negatively related to well-being (Franklin et al., 2013; Killingsworth & Gilbert, 2010). Hence, this presents evidence that high levels of salience can present an interference and thus distract from the current moment or task.

Surprisingly, the positivity of thoughts about mediated interactions was independent of their frequency (i.e., salience). We expected the effect of these thoughts to become weaker the more positive they were. The lack of an interaction suggests that the frequency of thoughts about mediated interactions is negatively related to well-being. Yet when thoughts are positive, this positivity may more than compensate the possible negative effect of frequency, yielding an overall positive effect of thoughts with positive valence. In other words, the frequency of thoughts may not play a role, as long as they are of positive valence. This has important implications for the online vigilance construct. Specifically, the pattern of effects provide first evidence for the theoretical proposition that online vigilance is not uniformly detrimental to well-being (Reinecke, 2018): The internalization of connectedness appears problematic if expressed in frequent distracting thoughts about

the online sphere, but it is by far more important whether these thoughts are positive or negative.

Interestingly, this conclusion does not hold up for thoughts about face-to-face interactions. Although they were of higher frequency than thoughts about mediated interactions, they were virtually unrelated to affective well-being. Like thoughts about mediated interactions, however, the valence of thoughts about face-to-face interactions was a strong positive predictor of well-being. That is, the more positive thoughts were about interactions that people had face-to-face in the last half an hour, the better they felt at the moment. Importantly, this positive relation was about as strong as that of the valence of thoughts about mediated interactions with well-being. These findings have several implications. First, thoughts about the online sphere might be more bothersome than thoughts about the offline sphere, possibly because they increase the salience of social pressure for constant connectedness (Halfmann & Rieger, 2019; Reinecke, Aufenanger, et al., 2017). Second, the valence of thoughts about mediated interactions can be just as beneficial for well-being as the valence of thoughts about face-to-face interactions, further supporting the argument that online vigilance can have both positive and negative effects on well-being. However, the analysis of thoughts about face-to-face interactions was purely exploratory and needs to be independently replicated before we can make strong conclusions.

Monitoring was negatively related to well-being. The more time people spent on social apps in the half hour before they answered a survey, the worse they felt at that moment. However, this effect was small and not robust. The robustness of the effect depended on analytical choices (e.g., outlier exclusion), similar to studies showing that the relation between screen time and media use can depend on analysis choices (e.g., Orben et al., 2019). Even if an effect exists, it is doubtful whether it is large enough to have practical consequences. As such, this component of online vigilance does not appear to be negatively related to well-being.

A similar argument applies to reactivity. Contrary to our prediction, reactivity was not significantly related to well-being and displayed a small effect. In other words, how quickly participants responded to surveys was not related to their well-being. This lack of a meaningful relation could have several reasons. First, it might indicate that the reactivity dimension is not associated with task interruptions. Alternatively, it might be associated with task interruptions, but these interruptions may not be severe enough to lead to interference or higher communication load and subsequently impair well-being. Both accounts would be in contrast to previous literature showing such interruptions can result in increased stress and pressure (Halfmann & Rieger, 2019; Reinecke, Aufenanger, et al., 2017).

Self-reported indicators of online vigilance correlated moderately with the online vigilance trait, in line with previous work (Reinecke et al., 2018). The self-reported indicators also displayed high correlations among each other. In contrast, the correlations between the behavioral indicators and the online vigilance trait were much lower. This pattern of

results has several implications. First, the lack of a relation between trait and behavior supports the view that trait measures might sometimes not be predictive of actual behavior (Masur, 2018). Similarly, people who report high levels of general awareness of online streams of communication might not express behavior in line with that self-assessment. This possibility does not invalidate online vigilance as a construct, but rather speaks to the larger issue of how predictive such person-level media-related variables are of actual behavior on the situation level (Ellis et al., 2019; Scharnow, 2016). Second, even when we assume the self-reported items are a better indicator of online vigilance, our exploratory analysis showed that these dimensions were not related to well-being. Hence, regardless of whether predictors were behavioral or self-reported, monitoring and reactivity were not meaningfully related to well-being.

Our study comes with several limitations. First, the lack of an interaction effect between salience and valence of mediated thoughts could reflect a lack of power. Because participants did not report salience to occur in every episode, we had to rely on a subset of the data, which greatly reduced observations per participant. However, inspecting the credible interval of the interaction effect indicates that the effect likely is small, possibly too small to have practical relevance. Future research should consider testing these effects with larger samples or more measurements per participant.

Second, it is unclear whether time spent with social apps is truly reflecting the monitoring dimension of online vigilance. We originally operationalized monitoring as phone checks unprompted by notifications. We assumed such phone checks were mostly non-purposeful because they did not serve an explicit goal (Hintze et al., 2017; Oulasvirta et al., 2012), thereby representing a manifestation of monitoring. Because the logging app did not consistently record notifications, we had to follow our preregistered decision tree and rely on time spent on social apps as a measure. Whereas someone high in monitoring would certainly spend more time on social apps, such a measure subsumes both use triggered by a notification and checks unprompted by notifications. That being said, there was a moderate correlation between self-reported monitoring and behavioral monitoring. Consequently, we believe time spent with social apps presents a suitable, but not ideal measure of the monitoring dimension. Because logging smartphone use is a complicated endeavor, future research should consider collaborations with scholars from other fields, such as computer science, to log notifications more reliably.

Third, it is likely that operationalizing reactivity as how quickly participants responded to surveys captures other processes in addition to reactivity. For example, such response time might reflect compliance to a considerable degree. Alternatively, participants received a higher compensation the more surveys they answered, possibly motivating them to respond faster. At the same time, someone high in reactivity will open notifications faster than someone low in reactivity, including surveys from an experience sampling app. However, there was only a low correlation between self-reported and behavioral reactivity. On the one hand, this might indicate that our behavioral reactivity measure was not adequate to capture this dimension of online vigilance. On the other hand, this

mismatch might merely reflect a disconnect between people's reported and their actual behavior (e.g., Ellis et al., 2019). It thus remains unclear whether the measure captures the whole spectrum of the reactivity dimension. We can only repeat the call for more research investigating the role of notifications.

Last, online vigilance on the trait level was only weakly related to trait well-being, in line with previous work demonstrating that direct relations were masked by mediators (Johannes et al., 2018). At least on the trait level, online vigilance may thus not directly relate to well-being, but rather through absentmindedness or stress (Reinecke et al., 2018). It is possible such mediating mechanisms have to be taken into account when investigating online vigilance and well-being on the state level. This requires both theoretical and methodological advances. Theoretically, there is still a vast avenue to develop theoretical models of how online vigilance relates to other constructs. Methodologically, there is a need for sophisticated analysis techniques that test larger theoretical models while taking nested data structures into account.

Overall, our study makes both theoretical and methodological contributions that enhance our understanding of the construct of online vigilance. We take into account that connectedness can take the form of an internalized psychological state that is expressed both cognitively and behaviorally. As such, we go beyond the simplistic view of screen time as the most important indicator of connectedness. Furthermore, we extend previous research by testing the expression of online vigilance and its relation to well-being on the situational level, employing adequate measures of both behavior and cognition. This approach allows us to understand online vigilance and its relation to well-being within a highly ecological setting. Methodologically, we preregistered the entire project, contributing to a more reliable knowledge base on media effects. Most important, our findings on salience and its valence demonstrate the need for nuance when studying media effects. Specifically, the results warrant caution not to treat cognitive preoccupation with mediated interactions as uniformly negative. Constant connectedness has become the norm among many users. More high-powered, preregistered research with refined measures is needed to better understand our constantly connected everyday lives.

Hard to Resist? The Effect of Smartphone Visibility and Notifications on Response Inhibition

This chapter is in press as Johannes, N., Veling, H., Verwijmeren, T., & Buijzen, M. (in press). Hard to resist? The effect of smartphone visibility and notifications on response inhibition. *Journal of Media Psychology*, Advance Online Publication. <https://doi.org/10.1027/1864-1105/a000248>



ABSTRACT

Because more and more young people are constantly presented with the opportunity to access information and connect to others via their smartphones, they report to be in a state of permanent alertness. In the current study, we define such a state as smartphone vigilance, an awareness that one can always get connected to others in combination with a permanent readiness to respond to incoming smartphone notifications. We hypothesized that constantly resisting the urge to interact with their phones draws on response inhibition, and hence interferes with students' ability to inhibit prepotent responses in a concurrent task. To test this, we conducted a preregistered experiment, employing a Bayesian sequential sampling design, where we manipulated smartphone visibility and smartphone notifications during a stop-signal task that measures the ability to inhibit prepotent responses. The task was constructed such that we could disentangle response inhibition from action selection. Results show that the mere visibility of a smartphone is sufficient to experience vigilance and distraction, and that this is enhanced when students receive notifications. Curiously enough, these strong experiences were unrelated to stop-signal task performance. These findings raise new questions about when and how smartphones can impact performance.

Every year, the publisher *Langenscheidt* elects the youth word of the year in Germany, which is supposed to best reflect generational trends and issues important to youth. For 2015, they chose “smombie”, a portmanteau of smartphone and zombie (*Süddeutsche Zeitung*, 2015). According to the publisher, the term represents a phenomenon that recent research has termed being “permanently online” (Vorderer & Kohring, 2013). More and more young users report being in a state of mind of permanent readiness to respond to their smartphones (Pew Research Center, 2015). Such a state of mind can be understood as vigilance (Bayer et al., 2015), because young users are constantly confronted with the desire to check their phones to satisfy social, informational, and hedonic needs (Chun, Lee, & Kim, 2012; Karapanos et al., 2016). Resisting those urges requires one of the most basic forms of executive functions (EF), namely behavioral inhibition (Friedman et al., 2008; Jurado & Rosselli, 2007). Therefore, with the current study we aimed to test whether smartphone vigilance indeed draws on inhibitory capacities, thereby decreasing performance in a simultaneous inhibition task.

SMARTPHONES AND INHIBITION FAILURE

Students in particular frequently fail to resist checking their smartphones, because smartphone checking has shown to be the most frequent interruption during self-study (Calderwood et al., 2014), characterized by nonpurposeful, reward-based checks (Oulasvirta et al., 2012). Giving in to those checks has been associated negatively with various performance outcomes, such as grades (for reviews see Chein et al., 2017; Q. Chen & Yan, 2016; van der Schuur et al., 2015).

Despite displaying a willingness to refrain from checking their phones in the face of more important tasks, students' repeated failures to do so demonstrates that inhibiting phone checking can be difficult. Inhibiting to give in to smartphone distractions can be considered an EF process which lies at the heart of self-regulation (Hofmann, Schmeichel, & Baddeley, 2012; Inzlicht & Berkman, 2015). EFs involve “a set of general-purpose control mechanisms, often linked to the pre-frontal cortex of the brain, that regulate the dynamics of human cognition and action” (Miyake & Friedman, 2012, p. 8). EF failure is related to various undesirable outcomes, such as overeating (Cserjési, Luminet, Poncelet, & Lénárd, 2009) and academic problems (J. R. Best, Miller, & Naglieri, 2011). Different authors have suggested different classifications for EFs, but there seems to be consensus on three components (e.g., Diamond, 2014): attentional control or working memory capacity; shifting, which refers to switching between tasks or mental sets; and inhibition, which refers to the suppression of task-irrelevant thoughts, actions, dominant responses, or urges.

Resisting thoughts about desirable outcomes provided by smartphones, and resisting the urge to check one's phone, lie at the intersection of two forms of inhibition: inhibitory control of attention and the inhibition of learned motor responses. The former describes deliberately suppressing attention to stimuli (Diamond, 2014), for instance smartphone notifications, that are irrelevant to the task at hand. The latter refers to inhibiting the

learned response to pick one's phone up either to check whether new notifications have come in or to respond to a notification (Soror, Hammer, Steelman, Davis, & Limayem, 2015). There are a number of empirical reasons to suspect that inhibition (failure) is indeed related to smartphone use. First, individuals with problematic smartphone (Roberts & Pirog, 2013; Smetaniuk, 2014) and instant messaging use (Levine, Waite, & Bowman, 2013) also display high impulsivity, as do those who engage in more multitasking (Sanbonmatsu, Strayer, Medeiros-Ward, & Watson, 2013). Second, Hadlington (2015) found a strong relationship between problematic mobile phone use and everyday cognitive failures. Finally, Sanbonmatsu et al. (2013) found that those with low executive control were more likely to report higher levels of media multitasking, including smartphone use.

SMARTPHONE VIGILANCE

With more and more users reporting to be in a state of alertness to respond to their devices (Pew Research Center, 2015) or even entrapment (Hall & Baym, 2012), we can understand this permanent alertness as a state of vigilance. Vigilance is traditionally defined in the context of monitoring work objectives as "the ability of organisms to maintain their focus of attention and to remain alert to stimuli over prolonged periods of time" (Warm et al., 2008, p. 433). Smartphone vigilance can be understood similarly, but not as the primary object of focused attention, but rather as ongoing alertness parallel to other tasks. Seo, Kim, and David (2015) refer to the state as connectedness, "the inclination or investment to remain connected with others or being available to others through phone and other mobile technologies" (p. 671). In the present study we define smartphone vigilance as a state of being aware that one can always get connected with others or access information, accompanied by a permanent readiness to respond to incoming smartphone stimuli (Bayer et al., 2015).

Smartphone vigilance may thus continuously interfere with the inhibition process. Students need to inhibit both their behaviorally learned response to check for notifications as well as thoughts about the potential rewards their phones offer. By taxing the EF of inhibition, smartphone vigilance may thus impair performance on a simultaneous task requiring inhibition.

Although such a position has not been explicitly tested, other research suggests that smartphone vigilance indeed taxes EFs. For instance, Shelton, Elliott, Eaves, and Exner (2009) showed that hearing a phone ring during a lexical decision task resulted in performance decrements; moreover, participants displayed increased recovery times from phone rings compared to other tones. In an innovative experiment, Stothart, Mitchum, and Yehmert (2015) demonstrated that students could be distracted by texts or calls without even responding to those notifications. During a sustained attention to response task (SART), participants received notifications on their own phones sent by the experimenters (not knowing they would be texted or called). During a SART, participants are to press a key every time a target number appears (1-9), unless a nontarget number is displayed (e.g., 3). Due to its repetitive nature, the SART requires a prolonged period

of attention. Receiving notifications resulted in diminished performance in the SART. Further, the authors concluded that the magnitude of the effect was comparable to using the phone whilst driving. Thornton, Faires, Robbins, and Rollins (2014) employed a similar design, but showed that receiving a notification might not be necessary. In two studies, students performed attention tasks with a phone on the table that did not receive any notifications. Just the mere presence of the experimenter's or participants' phones led to diminished performance.

However, none of these studies explicitly manipulated or measured vigilance, a psychological state. Rather, they manipulated smartphone visibility and notifications, which we believe result in vigilance, because they best mimic real-life situations that induce vigilance. In addition, because both Stothart et al. (2015) and Thornton et al. (2014) measured attention, the effect of smartphone vigilance on inhibition remains to be examined. Consequently, we planned to extend previous research using similar manipulations (visibility and notifications) which should induce smartphone vigilance. Besides measuring self-reported vigilance, we tested whether these manipulations influence performance on a validated and established task tapping inhibition, a modified stop-signal task (Logan, 1994; Verbruggen & Logan, 2008b).

STOP-SIGNAL TASK

In a stop-signal task, participants react to a stimulus during go-trials by, for instance, indicating the direction of arrows, unless they hear or see a stop signal (i.e., stop-signal trials). Timing of the stop signal is adjusted dynamically depending on participants' performance using a staircase procedure: The stop signal is presented earlier when participants fail to stop and later when participants succeed in stopping. This procedure allows for estimating the time to stop a response (the so-called stop signal reaction time, SSRT), which is considered a measure of response inhibition. We expected smartphone visibility and notifications to induce vigilance and thus impair response inhibition, thereby increasing SSRT.

Importantly, recent work suggests that stop-signal trials not only require response inhibition; participants also have to update their current action plan (i.e., update the automated go-response to the alternative response, stopping) as well as update their attention (i.e., detecting the stop-signal; Verbruggen et al., 2010; Verbruggen, Stevens, et al., 2014). For the sake of brevity, we refer to these processes as action selection. Therefore, to be able to distinguish between response inhibition and action selection, we also employed trials that require a double-response. During double-response trials, participants not only carry out the automated response, but also a second one (e.g., pressing the spacebar after categorizing the direction of an arrow). The task thus permitted us to explore the possibility that smartphone visibility and notifications have an effect on action selection (i.e., double-response reaction times, DTR2). We hypothesized that compared to a no-visibility-no-notifications control condition, the visibility-without-notifications-condition would have a negative effect on response inhibition (H_{1a}), and the visibility-with-notifications-condition

would have a negative effect on response inhibition as well (H_{1b}). In addition, we explored whether visibility without notifications differed from visibility with notifications.

METHOD

We conducted an experiment in order to examine the influence of smartphone visibility and notifications on the EF of inhibition by employing a modified stop-signal task, that is, a context-cueing task. Readers can find study materials and data on the Open Science Framework (<https://osf.io/k3p54/>). We obtained approval from the institute's IRB (approval code: ECSW2016-0905-392a).

PARTICIPANTS AND SAMPLING DESIGN

Even though previous research on smartphone vigilance found medium-sized effects (Stothart et al., 2015; Thornton et al., 2014), these effects were demonstrated mainly for attention, not inhibition. Consequently, we could not be certain those effect sizes would apply to our experiment. Without certainty about an expected effect size, a power analysis for a frequentist analysis would likely yield an inaccurate sample size estimation. In addition, preregistered reports should allow to quantify support in the data for possible null-findings, which is not possible under a frequentist framework (Wagenmakers, 2007). Therefore, we employed a sequential Bayesian sampling design (Schönbrodt, Wagenmakers, Zehetleitner, & Perugini, 2017). Such a design allowed us to address both issues of power and support for null effects: First, sequential Bayesian designs are flexible and resource-efficient, because they allowed us to continuously monitor the data, thus circumventing the risk of incorrectly calculating power. Second, Bayesian analyses can quantify support for the lack of an effect (Schönbrodt et al., 2017).

Bayesian analyses let researchers assign a probability distribution of effect sizes that they assume is plausible for their study. This so-called prior distribution is then compared to the likelihood distribution for the observed data to form the posterior distribution. Thus, the posterior distribution represents the distribution of effect sizes for the observed data taking into account the researcher's prior belief about the effect. Comparing the posterior with the prior tells us how much the information from the data has updated our prior belief. In addition to estimating the posterior distribution of effect sizes (i.e., parameter estimation), Bayesian analyses can also be used to select and compare competing hypotheses by using the so-called Bayes Factor (BF). Comparing the prior to the posterior at an effect size of zero, the BF indicates how much more the data are likely under the alternative hypothesis than under the null hypothesis (or vice versa). For instance, $BF_{10} = 6$ means the data are six times more likely under the alternative hypothesis than under the null hypothesis, and $BF_{01} = 6$ means the data are six times more likely under the null hypothesis than under the alternative hypothesis (for an introduction see Wagenmakers et al., 2018).

We set a minimum sample size of $n = 20$ per condition and a maximum sample size of $n = 50$ (after implementing exclusions). As a stopping rule, we set an a priori threshold of 6 for the BF for BF_{10} and 6 for BF_{01} as recommended by Schönbrodt et al. (2017) for all direct comparisons. Because we did not reach the boundary conditions for H_{1a} or H_{1b} after the minimum sample size, we continued sampling until the maximum sample size. Overall, 178 people participated in our study, of which we retained 154 valid cases after applying exclusion criteria (see below). Participants (113 female, 73%, $M_{age} = 21.70$, $SD_{age} = 2.58$) were students from a university in the Netherlands, who received €5 or course credit. They owned a smartphone for multiple years ($M = 6.74$, $SD = 2.04$), and most of them estimated to check their phones rather frequently per day, with 84% indicating to check their phones 20 times or more.

PROCEDURE

MANIPULATION

We employed a between-subjects design with three groups (no-visibility-no-notifications vs. visibility-without-notifications vs. visibility-with-notifications). Before signing up for the study, participants were informed that the experiment would be about cognitive performance and smartphones, and that they should be willing to have an experimenter change their phone settings to silent or vibrate mode.

Upon arrival, the experimenter welcomed participants and randomly assigned them to one of the conditions, followed by the context-cueing task. The experimenter told participants that they would set their phones to either silent or vibrate mode. By setting participants' phones either to flight mode (visibility-without-notifications condition, the alleged silent condition) or disconnected from the internet, with vibrate mode on (visibility-with-notifications condition, the alleged vibrate condition), we planned to induce vigilance. In the no-visibility-no-notifications control condition, a notebook was placed on the table.

In the visibility condition, participants' phones were set to flight mode with silent mode on to make sure no notifications would come in. Yet, participants did not know whether their phone was in silent or in vibrate mode; thus, because they believed a notification could come in at any time, they were likely to be vigilant. The manipulation thus aimed to reproduce the vigilant status of a majority of students in their everyday lives, but without the possibility of actually receiving a notification.

In the notifications condition, participants' phones were disconnected from the internet, with vibrate mode on. This way, they could only receive regular SMS, whose frequency is negligible compared to instant messaging services such as WhatsApp (bitkom, 2015). The chance was thus minimized that they received a notification not sent by the experimenter (preventing exclusion; see below). During the last 10 of 32 practice trials of each block, they received three text messages, making their phones vibrate, separated by seven seconds so participants would not mistake the notifications for a call. The SMS were sent by the program to the number they indicated when registering for the experiment. Because

participants were not allowed to check their phones, they could not be sure whether one of their personal contacts had messaged them or the experimenter. The manipulation thus aimed to induce the status of alertness to check one's messages, that is, participants were likely to be vigilant.

In the no-visibility-no-notifications control condition, participants set their phones to silent mode and stored it in the pockets of their jackets or in their handbags, and put them in the corner of the cubicle. This way, participants could not feel their phones in their pockets, limiting the potentially confounding effect of sensory perception of their smartphone, which in itself could induce vigilance. We decided not to remove participants' phones from the room, because smartphone separation has shown to be detrimental to EF in its own right (Hartanto & Yang, 2016). In addition, a notebook of similar size to a smartphone was placed on the table to make sure differences between the groups did not arise simply because there was a graspable object on the table (Przybylski & Weinstein, 2013).

The notebook or participants' phones were put on the table next to their dominant hand, display-down, so participants were blind to condition and could not see whether their displays lighted up. After the experiment, participants were fully debriefed and received their compensation.

RESPONSE INHIBITION MEASURE

To measure response inhibition, we employed a version of the stop-signal task known as the context-cueing task (e.g., Verbruggen et al., 2010). This task allows a distinction between possible effects of the manipulations on response inhibition versus action selection. During this task, participants were instructed to categorize the direction of an arrow as left or right by pressing the 'U' or 'I' key on the keyboard. Participants were first presented with a shape for 500ms within which the arrow appeared. The shape served as context cue. Participants completed two blocks of trials: In one block the shape was a circle cueing stop trials; in the second block the shape was a square cueing double-response trials. The order of blocks was counterbalanced. On 30% of trials, the context cue turned bold (signal trial) for 250ms. For stop-signals, participants had to inhibit their response to categorize the arrow. For double-response trials, participants had to categorize the arrow and additionally press the space bar. The arrow disappeared after participants made their choice or after 1500ms. Inter-trial interval was 250ms. For double-response trials, the frame randomly turned bold after a delay of 100, 250, or 400ms (SOA). SOA for stop-signal trials followed a staircase tracking procedure (Verbruggen, Chambers, & Logan, 2013): SOA started at 250ms; when inhibition was successful, SOA increased by 50ms; when inhibition was unsuccessful, SOA decreased by 50ms.

We instructed participants that sometimes it would be impossible to be successful on stop-signal trials, but that they should not wait for the shape to turn bold and categorize as quickly and accurately as possible. Each block consisted of 120 trials, each preceded by 32 practice trials. During the last 10 practice trials of each block, the notification

condition received three notifications. The dependent measures were SSRT (a measure of response inhibition) and DRT2 (a measure of action selection). Because the probability of responding on stop-signal trials was not .50 ($M = .53$, $SD = .11$; range: .28-.89), as assumed by the mean estimation method, we calculated SSRT calculated by subtracting the average SOA ($M = 218.3$, $SD = 166.6$, range: 25.0-822.2) from the finishing time of the stop process (integration method; see Verbruggen et al., 2013).

The context-cueing task and notification program were coded in Python, Version 2.7; notifications were sent using Twilio, a cloud-based interface to send text messages, following the code as provided by Stothart et al. (2015).

MANIPULATION CHECKS

With visibility and notifications we did not directly manipulate the psychological state of vigilance. Therefore, we had to ensure these manipulations did, in fact, lead to vigilance. After the context cueing task, participants answered nine items about their smartphone vigilance during the task (e.g., "My smartphone occupied my thoughts, even though I was doing the task") on a scale ranging from 1 (*strongly disagree*) to 5 (*strongly agree*) that we adapted from online vigilance trait items by Reinecke et al. (2017). If independent Bayesian t-tests showed participants in the visibility and notifications conditions to be more vigilant than the control condition, we could be more confident that any effects on inhibition were indeed caused by smartphone vigilance, as posited in our hypotheses. Further, participants indicated whether their phone vibrated during the experiment, how distracting their phone was, whether their phone was in their line of sight during the experiment, and whether they touched their phone.

ADDITIONAL MEASURES

In addition to demographic information, we assessed several personality traits that have shown to be related to smartphone use in order to describe the population and for exploratory analyses only: fear of missing out (Przybylski, Murayama, DeHaan, & Gladwell, 2013), susceptibility to boredom (Mercer-Lynn, Flora, Fahlman, & Eastwood, 2013), need to belong (Leary, Kelly, Cottrell, & Schreindorfer, 2007), and need for popularity (Santor, Messervey, & Kusumakar, 2000). We will not report on those exploratory measures. They are available on the OSF.

EXCLUSION CRITERIA

We applied the following a priori exclusion criteria before data analyses. We did not allow participants to touch their phones in order to alleviate the presumed state of vigilance; to control such phone interaction, all sessions were recorded with webcams in the cubicle. Consequently, we excluded five participants who touched their phones during the experiment. In addition, we excluded ten participants in the notification condition, because their phone did not vibrate even though they received the text messages, due to customized notification profiles that the experimenter could not access. Further, we

excluded two participants who received SMS that were not sent by us. Following our exclusion criteria on the participant level for the context cueing task, we excluded one participant who had a negative SSRT. Last, because categorizing the arrow was fairly easy, we excluded six participants with lower accuracy than 90% on go-trials. After applying exclusions, we slightly exceeded our maximum sample size of 50 per group (control: $n = 51$; visibility: $n = 53$; notification: $n = 50$), due to our randomization procedure. On the trial level, we excluded RTs 3 SD above or below the respective mean on go-trials on the SST (1.11%) and DRT (1.59%), as well as on the second response of double-response trials (.67%). Last, we excluded RTs in go-trials below 200ms (.86%).

RESULTS

We conducted all Bayesian analyses with JASP (JASP Team, 2017). All t-tests are two-sided unless otherwise specified, with the standard JASP Cauchy prior. Within JASP, we also conducted robustness checks with different priors. A robustness check is used to examine how the BF changes if one has a different prior beliefs about the effect size (Wagenmakers et al., 2018).

PREREGISTERED ANALYSES

MANIPULATION CHECKS

As manipulation checks, we asked participants whether their phone was in their line of sight throughout the entire experiment. As expected, all participants in the control condition said their phone was not in their line of sight. The majority of participants in the visibility condition (81%) and the notification condition (86%) indicated the same. Even though from our video recordings it was clear that phones were right next to participants' dominant hand, apparently some participants defined their line of sight strictly as the monitor.

In addition, we asked participants whether their phone vibrated during the experiment. As expected, nobody in the control condition perceived a vibration; similarly, nobody except one participant perceived a vibration in the visibility condition. We checked the participant's video again, but could not hear a vibration. Furthermore, everybody in the notification condition perceived the vibration. Last, nobody of the final sample indicated that they touched their phone.

Further, when asked on a visual analogue scale ranging from 0 to 100 how distracting the phone was during the task, we found the expected pattern, such that those in the control group reported close to no distraction ($M = 1.29$, $SD = 8.56$), those in the visibility condition minimum distraction ($M = 10.53$, $SD = 17.32$), and those in the notification condition considerable distraction ($M = 43.10$, $SD = 24.68$). All of those differences among groups were more plausible under the alternative hypothesis than under the null model

(all $BF_{10} > 32$, $d > .67$), with medium to large effect sizes, suggesting participants indeed perceived our manipulation as distracting (see Table 1).

Table 1. Bayes Factors for all hypothesized group comparisons

Variable	Groups Compared					
	Control-Visibility		Control-Notification		Visibility-Notification	
	BF_{10}	BF_{01}	BF_{10}	BF_{01}	BF_{10}	BF_{01}
Distraction	32.65	.03	3.61e+16	2.77e-17	1.06e+9	9.43e-10
Vigilance	37.50	.03	3.14e+7	3.18e-8	97.11	.01
SSRT	.43	2.32	.21	4.76	.42	2.36
DRT2	.21	4.81	.21	4.76	.21	4.80

Finally, we tested whether our manipulations induced vigilance. Overall, vigilance was below the midpoint of the five-point scale, and skewed towards the lower end ($M = 1.60$, $SD = .66$). As expected, vigilance was lowest in the control condition ($M = 1.22$, $SD = .44$), followed by the visibility condition ($M = 1.56$, $SD = .56$) and the notification condition ($M = 2.03$, $SD = .69$). Bayesian independent t-tests indicated that the data were more likely under the alternative hypothesis of a difference between groups than under the null hypothesis: control-visibility: $BF_{10} = 37.50$, 95% CI: [-1.03, -.24], $d = -.68$; control-notification: $BF_{10} = 3.14e+7$, 95% CI: [-1.78, -.91], $d = -1.40$; visibility-notification: $BF_{10} = 97.11$, 95% CI: [.31, 1.08], $d = .75$. All Bayes Factors displayed strong to extreme evidence (Lee & Wagenmakers, 2013) with medium to large effect sizes in favor of the alternative hypothesis (see Table 1). Overall, our manipulation checks showed that our manipulation was successful, with participants being aware of their phones when in the visibility or notification condition, perceiving it as more distracting, and experiencing more vigilance. Crucially, participants in the notification condition did receive and notice the text messages we sent.

CONFIRMATORY ANALYSES CONTEXT-CUING TASK

Inspecting the context-cueing task, participants almost never missed the second response in the DRT (>99%), and overall accuracy for all conditions was almost identical (all > 97%). Overall SSRT ($M = 226.8$, $SD = 59.8$, range: 20.2 – 337.1) was in a similar range as previous work (e.g., Verbruggen & Logan, 2009). As expected, overall DRT2 was much higher ($M = 729.3$, $SD = 80.4$, range: 584.4 – 1039.2).

Next, we tested our hypotheses on task performance (see Table 1). H_{1a} stated that the visibility condition would result in higher SSRT than the control condition. Contrary to our expectations, SSRT was comparable in the control condition ($M = 232.1$, $SD = 51.3$) and visibility condition ($M = .216.5$, $SD = 70.8$), with the Bayes Factor indicating that the data were about twice ($BF_{01} = 2.32$, 95% CI: [-.15, .60], $d = .25$) as likely under the null hypothesis than under the alternative hypothesis, which qualifies as inconclusive evidence (Lee &

Wagenmakers, 2013). Inspecting Figure 1 (upper panel), the robustness check shows that even widening the prior distribution to assign less mass to a null effect does not substantially increase the Bayes Factor. In addition, the sequential analysis perpetuates the inconclusive nature of the finding, as the Bayes Factor hovers around 1, indicating the data are equally likely under the null and the alternative hypothesis.

H_{1b} stated that the notification condition would result in higher SSRT than the control condition. Surprisingly, SSRT in the notification condition ($M = 232.3$, $SD = 54.8$) was almost identical with that in the control condition; indeed, the data were about five times ($BF_{01} = 4.76$, 95% CI: [-.37, .36], $d = -.004$) more likely under a model assuming no difference between the conditions than under the alternative hypothesis, which qualifies as moderate evidence. The robustness check in Figure 1 (middle panel) displays an increase in evidence for the null hypothesis as less mass is assigned to zero. Likewise, the sequential analysis supports evidence for the null hypothesis, as the Bayes Factor continually increases.

Last, we also explored whether there were differences in SSRT between the visibility and notification condition. Again, the difference between the two conditions was inconclusive; the data were about twice as likely under the null hypothesis as under the alternative hypothesis of an effect ($BF_{01} = 2.36$, 95% CI: [-.14, .60], $d = .25$). As Figure 1 (lower panel) shows, the finding can be classified as inconclusive, as even giving less prior plausibility to the null hypothesis does not increase support for the null substantially. Just as with H_{1a} , the sequential analysis displays Bayes Factors that hover around 1. Because smartphone vigilance might not have an effect on response inhibition, but rather on action selection, we investigated whether there were differences between the conditions on DRT2. Comparing the control condition ($M = 728.8$, $SD = 80.4$) with the visibility condition ($M = 729.9$, $SD = 73.5$) showed moderate support for the null hypothesis ($BF_{01} = 4.81$, 95% CI: [-.38, .35], $d = -.02$) which steadily increased with wider prior distributions and a clear trend in the Bayes Factors towards H_0 for the sequential analysis.

Similarly, DRT2 between control condition and notification condition ($M = 729.2$, $SD = 88.6$) were extremely similar, again with moderate support for the null hypothesis of no difference ($BF_{01} = 4.76$, 95% CI: [-.36, .37], $d = -.05$). Widening the prior did again increase support for the null hypothesis; the sequential analysis displayed a trend towards H_0 as well.

Last, comparing the visibility and notification conditions, the data were about five times ($BF_{01} = 4.80$, 95% CI: [-.37, .35], $d = -.01$) more likely under the null hypothesis of no difference, indicating moderate support. As before, increasing the prior width increased support for H_0 , and the sequential analysis demonstrated a clear trend towards H_0 as well.

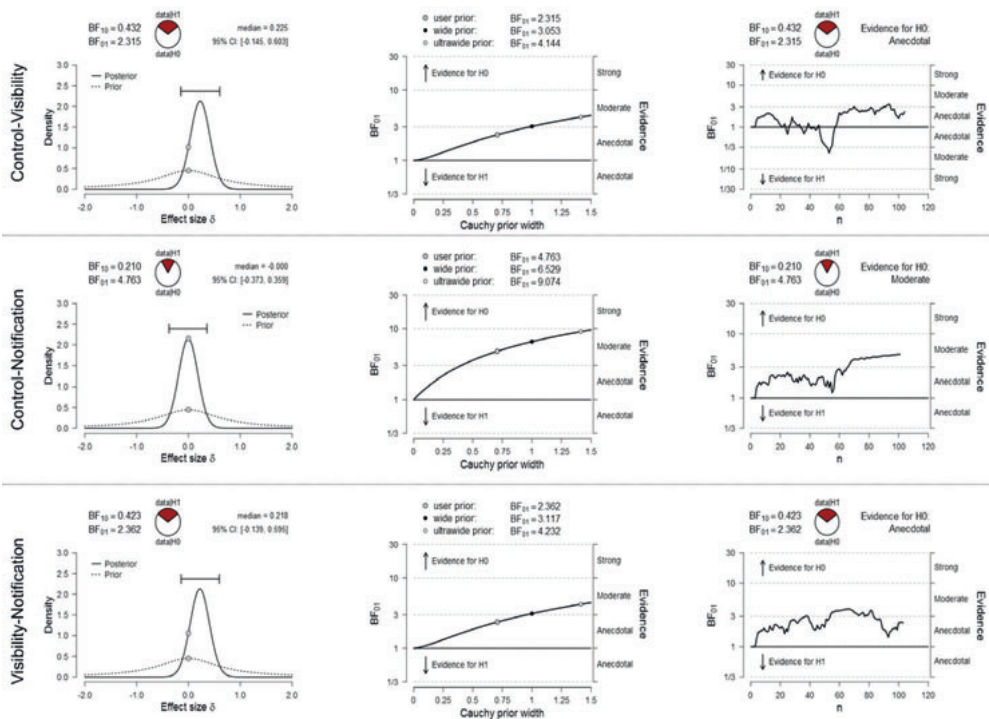


Figure 1. Prior and posterior distribution (left), Bayes Factor robustness check (middle), and sequential analysis (right) for all independent Bayesian t-tests on SSRT. Upper panel compares control condition with visibility condition. Middle panel compares control with notification. Lower panel compares visibility with notification.

EXPLORATORY ANALYSES CONTEXT CUEING-TASK

In accordance with our preregistration, we investigated an alternative explanation should we not find the expected effects, namely, proactive control. Previous work has shown that GoRTs are slower in stop-signal blocks compared to double response blocks (e.g., Verbruggen & Logan, 2009). This effect is assumed to reflect proactive control in order to avoid failing to stop in time. Indeed, and consistent with this work, GoRTs were higher in the SST ($M = 466.9$, $SD = 168.8$) compared to the DRT blocks (335.7 , $SD = 43.3$; Bayesian paired-sample t-test: $BF_{10} = 2.64e+15$, 95% CI: [.61, .98], $d = .81$), reflected on the overall positive difference score ($M = 131.1$, $SD = 162.7$).

Similarly, we might expect enhanced proactive control in the notification condition compared to the control condition: When participants proactively control their responses toward their phone, they should become slower on GoRTs even when the task does not require control, because the proactive control aimed at their phones also slows GoRTs.

This would mean that in the notification condition GoRTs in the double response block should be more similar to GoRTs in the stop block compared to this difference in the control condition.

To test this, we conducted independent Bayesian t-tests between the notification condition and the control condition on the difference scores between GoRTs in the stop-signal block and GoRTs in the double-response block (higher scores indicate more active proactive control). Crucially, we did not find any indication that our manipulations induced proactive control; instead, there was anecdotal to moderate evidence ($BF_{01} = 3.71$, 95% CI: [-.50, .24], $d = -.15$) that the data were more likely under the null hypothesis than under the alternative hypothesis (control: $M = 114.3$, $SD = 140.2$; notification: $M = 138.7$, $SD = 184.8$).

NONPREREGISTERED ANALYSES

Previous research (Stothart et al., 2015) sent notifications throughout the entire block, not just during the practice trials. To see whether we could replicate the effect, we compared accuracy on no-signal trials on the last ten trials of the practice block (i.e., when notifications came in) between the notification condition and the control condition. Even though accuracy during those trials was slightly lower in the notification condition ($M = 95.7$, $SD = 6.1$) than in the control condition ($M = 96.8$, $SD = 6.1$), the data were more likely under the null hypothesis ($BF_{01} = 3.27$, 95% CI: [-.21, .53], $d = .18$). This constitutes anecdotal to moderate evidence against a distracting effect of the notifications when they came in.

Last, to explore whether self-reported vigilance was related to response inhibition, we correlated vigilance scores with SSRT. The data did not support such a relationship; to the contrary, there was moderate evidence for the null model ($BF_{01} = 6.39$, 95% CI: [-.23, .08], $r = -.08$).

DISCUSSION

Contrary to our expectations, our experiment generally yielded results that support a lack of an effect of smartphone visibility and notifications on response inhibition. We can state with moderate certainty that there was no effect of the notification condition on response inhibition in our data. This lack of an effect was less pronounced when comparing the control and visibility condition, as the differences in inhibition were only slightly more likely assuming no effect. Similarly, the differences between visibility and notification provided only weak evidence for the lack of an effect. Although we collected our predefined maximum sample size, none of the Bayes Factors reached our stopping rule of $BF = 6$, so the data are not conclusive. Taken together, there is no evidence that smartphone vigilance had an effect on response inhibition; if anything, our study supports the lack of an effect.

In addition, our design allowed us to distinguish between response inhibition and action selection, which subsumes action plan updating and attention updating. Mirroring

the results for inhibition, it appeared that smartphone visibility and notifications did not have an effect on action selection either. Yet, as stated above, no Bayes Factor reached the threshold of six, so the overall evidence for the lack of an effect is only moderate.

Interestingly, effects emerged very clearly on self-reported measures: Just having a smartphone on the table increased self-reported vigilance, such as the urge to check their phone or their thoughts about what was going on with their phone. This vigilance was even higher when people received notifications, but could not check them. Similarly, visibility and notifications were perceived as very distracting. Consequently, participants were aware of their phones and they reported both a strong urge to check it as well as cognitive preoccupation with it.

MANIPULATION AND TASK

When comparing self-reported and behavioral data, participants indicated that they felt like smartphones were distracting and made them vigilant. However, our results imply they did not, in fact, affect executive functioning, neither response inhibition nor action selection. In our view, there are two possible methodological explanations why we did not find effects of smartphone vigilance on inhibition.

First, participants may not have been vigilant. That is, participants might have guessed that their phone was supposed to be distracting, and, consequently, reported higher levels of distraction and vigilance because they felt they were expected to, but did not experience this state. Although we cannot rule out such an explanation, participants were not aware of the other conditions; it seems unlikely that participants in the visibility condition reported much higher vigilance than the control condition, but lower than the notifications condition as a demand artifact. Moreover, if demand artifacts were an issue, it is not so clear why participants did not perform worse on the task (e.g., by making more errors), given that accuracy in the three conditions was close to identical. Furthermore, smartphone visibility and notifications were manipulated in a controlled and objective manner. Therefore, we believe it is reasonable to assume our manipulation did indeed induce vigilance.

Second, low validity of our inhibition measure could also account for our null-findings. However, all parameters we found are similar to that of previous research or follow logically from the premises of the task. First, overall SSRT and accuracy were within the range of what previous research found that used a relatively easy categorization task such as in our study (Verbruggen & Logan, 2009), and SSRT was slightly lower than previous research employing a more complicated task (Verbruggen et al., 2010). Second, because the DRT does not require stopping and, consequently, proactive control, GoRTs on the stop-signal task were higher on the SST than on the DRT, which adds to the validity of the measure (Verbruggen & Logan, 2009). Furthermore, the probability to respond on signal-trials as well as its range were comparable to previous work, which attests to the adequacy of the staircase procedure we employed (Verbruggen et al., 2010; Verbruggen & Logan, 2009).

Overall, then, it appears unlikely the measurement validity of the task was responsible for the lack of an effect.

RESOURCES, AUTOMATIZED VIGILANCE, AND PERSONALITY MODERATORS

In addition to methodological explanations, we believe there are theoretical accounts for why vigilance did not affect inhibition. First, the effect of smartphone vigilance on response inhibition could have been masked by increased performance due to increased recruitment of cognitive resources. Two related theoretical accounts could explain such enhanced control. From an avoidance cues account (Koch, Holland, & van Knippenberg, 2008), not being allowed to touch one's phone could have served as a cue to enter a state of alertness, recruiting more cognitive resources. Therefore, it is possible that the distraction of their phones interfered with inhibitory processes, while the additional resources recruited because of the avoidance mindset offset this interference. From the inhibitory spill-over account (Tuk, Zhang, & Sweldens, 2015), inhibitory capacity is not specific to one domain; rather, if one has to inhibit a response in one domain, it facilitates inhibition in an unrelated domain, as long as both processes happen simultaneously. As such, not giving in to the urge to check one's phone or think about it constitutes inhibition in one domain, which could facilitate response inhibition during the stop-signal task. Just as with avoidance cues, this spill-over could have counteracted against the interfering effect of smartphone vigilance. However, under both frameworks, the smartphone manipulations should have also facilitated proactive control, for which we found no evidence.

Second, at this point of smartphone saturation and constant connectivity, users may have grown accustomed to being vigilant at all times to a degree that it does not affect executive control anymore. Supporting such a position, we observed overall rather low levels of vigilance for the entire sample, below the midpoint of the scale, which could be an indication that participants were not overly vigilant. Instead, smartphone vigilance could have become automatized. As more recent work suggests, smartphone or online vigilance has likely become the norm among users (Klimmt et al., 2018). Evidence for such an assumption comes from Reinecke et al. (2017, May) who also reported means close to or below the midpoint of the scale. Thus, the vigilance we induced might have been strong enough to manifest itself on a self-reported level in the expected pattern, but it might be too automatized to affect behavior (Potter, 2011).

Third, expanding the point of automatized vigilance, it is likely that the effect of vigilance depends on personality factors. For example, users who have learnt to benefit from the constant social support their smartphones provide them might experience more intense vigilance which impedes performance (Reinecke, 2018). Similarly, users with a high fear of missing out may be particularly susceptible to smartphone cues (Przybylski et al., 2013). However, it is unclear whether such personality characteristics would raise the threshold of automatized vigilance or lower it. We invite researchers to use the data on personality traits we have collected and explore possible moderators. Moreover, there is a need for

research examining how much variation in vigilance there is between different users and how those users differ on other personality traits.

Finally, it is possible that vigilance influences certain executive control components such as sustained attention, but not others such as response inhibition (e.g., Stothart et al., 2015; Thornton et al., 2014). Future research may examine this possibility systematically.

IMPLICATIONS AND CONCLUSION

We investigated the assumption that smartphones interfere with response inhibition, using a large sample with an economic, flexible sampling design in a highly controlled laboratory experiment with a novel, innovative manipulation. We were unable to obtain evidence for the assumption that smartphones interfere with the inhibition process. Even though a lot of users complain that their phones put them in a state of alertness, which we termed smartphone vigilance, their phones did indeed make them feel vigilant, but did not interfere with executive functioning. Comparing our findings with research demonstrating a negative effect of smartphone use, for example in the form of multitasking (van der Schuur et al., 2015), shows that there is a need for subsequent studies investigating the difference between the mere presence and actual use of smartphones. As participants in our study were not allowed to touch their phones, restricting smartphone access appears to be beneficial to performance, which lends support to policies banning smartphone use in class. To conclude, our findings call for a better understanding under what conditions smartphones impair performance.

No Evidence that Smartphone Notifications Lead to Goal-Neglect

This chapter has been made publicly available as Johannes, N., Veling, H., & Buijzen, M. (2019). No evidence that smartphone notifications lead to goal-neglect [Preprint]. <https://doi.org/10.31234/osf.io/5me97>.



ABSTRACT

These days, young people report to be in a state of permanent alertness due to their smartphones. This state has been defined as smartphone vigilance, an awareness that one can always get connected to others in combination with a permanent readiness to respond to incoming smartphone notifications. We argue that receiving a notification makes users vigilant and activates goals (e.g., checking the message) that interfere with other goals needed to perform a task. We thus hypothesized that smartphone vigilance impairs maintenance of current task-goals in working memory, resulting in increased goal-neglect. To test this hypothesis, we conducted a preregistered experiment that examined the effect of smartphone vigilance (incoming notifications) on goal-neglect in a modified Stroop task. We found evidence that participants perceived notifications as distracting, but vigilance did not lead to increased goal-neglect. To the contrary, there was tentative evidence that vigilant participants performed better at the task.

As smartphone ownership approaches saturation in Western countries, especially younger users find themselves “permanently online” (Klimmt et al., 2018; Vorderer & Kohring, 2013). Interestingly, those users report to be alert at all times, expecting to satisfy informational, social, or hedonic needs through their smartphones (Bayer et al., 2018; Kanjo, Kuss, & Ang, 2017; Karapanos et al., 2016; Kardos et al., 2018). In a previous study, we defined this alertness as smartphone vigilance, “a state of being aware that one can always get connected with others or access information, accompanied by a permanent readiness to respond to incoming smartphone stimuli” (Johannes et al., in press). We predicted that constant preoccupation in the form of smartphone vigilance would interfere with the executive function of response inhibition (Diamond, 2014). Surprisingly, although we found a large effect of a smartphone vigilance manipulation on self-reported distraction and vigilance, there was no evidence for a negative effect on response inhibition.

These results suggest that smartphone vigilance does not influence the executive function of inhibition, which is in line with recent research demonstrating that smartphone exposure is not related to neurological change in cognition (Mohan, Khaliq, Panwar, & Vaney, 2016). Thus, inhibition impairment cannot explain the ample evidence that receiving notifications affects cognitive performance and attention (Shelton et al., 2009; Stothart et al., 2015; Thornton et al., 2014; Ward et al., 2017). Here, we propose that this negative effect might be explained by smartphone vigilance impairing another executive function, that of working memory. We argue that during vigilance, goals become activated, such as connecting to others (Kardos et al., 2018), that are competing with other goals in working memory, such as reading a paper (Cutino & Nees, 2016). In other words, once the phone vibrates, the notification activates the goal to pursue social or informational goals (i.e., checking the notification content), interfering with other goals kept in working memory (LaRose, 2015). Hence, we expected smartphone vigilance to interfere with the executive function of working memory capacity, signified by goal-neglect.

In order to assess goal-neglect, we employed the modified Stroop task as used by Kane and Engle (2003). During a Stroop task, participants have the goal to identify the ink color of a word. However, when word and ink are incongruent (e.g., RED in the color blue), the automatic tendency to read out the word presents interfering information, which requires participants to retrieve the goal of identifying the ink in order to resolve the conflict. Importantly, participants completed two versions of the Stroop task. First, during the majority-congruent version, 75% of trials are congruent (e.g., RED in the color red). No goal-maintenance is required during congruent trials. Only on occasional incongruent trials must they retrieve the goal of identifying the ink because of conflicting information. Thus, having that many congruent trials makes it harder to maintain the goal of identifying the ink color, which results in more errors during the occasional incongruent trials. Second, during the no-congruent version, 0% of trials are congruent. In this case, all trials are incongruent and require participants to constantly maintain the goal of identifying the ink color and not read the word. Thus, by repeatedly enacting the goal through their own behavior, goal-maintenance becomes easier for participants.

Consequently, in addition to the well-established Stroop effect (i.e., more errors on incongruent compared to neutral trials), the amount of goal-neglect shows in the interaction of Stroop effect and Stroop version. There should be no difference between the Stroop versions on the amount of errors on neutral trials. However, because goal-maintenance is harder on the majority-congruent Stroop version, there should be more errors on incongruent trials on the majority-congruent version compared to the no-congruent version. Crucially, to obtain the goal-neglect effect, Kane and Engle (2003) recommend this exact order of Stroop versions.

We predicted that smartphone vigilance would strengthen this goal-neglect effect in the vigilance condition compared to the control condition. Specifically, for smartphone vigilance to affect goal-neglect, we expected that there would be a larger difference in errors on incongruent trials in the vigilance condition compared to the control condition, as smartphone goals interfere with maintaining the goal to identify the ink.

METHOD

We preregistered our hypothesis, stopping rule, exclusion criteria, and analysis. The preregistration, data, analyses, and materials are on the Open Science Framework (<https://osf.io/z6aed/>). The study had IRB approval.

DESIGN

The experiment employed a 2 (smartphone vigilance: control, notifications) x 2 (Stroop version: no-congruent, majority-congruent) x 3 (trial type: neutral, incongruent, congruent) factorial design, with the first factor between subjects.

PROCEDURE AND MANIPULATION

The manipulation to induce the smartphone vigilance was identical to our previous study (Johannes et al., in press) and aimed to reproduce the feeling of vigilance that participants encounter in their everyday lives, namely, being in a state of alertness when receiving a notification.

In the notification condition, participants were informed that the experimenter would set their phones to either silent or vibrate, but participants were not allowed to check the setting of their phone. The experimenter then set the participant's phone to vibration mode, but also switched off Wi-Fi and mobile data services. Thus, only SMS could be received, which minimized the chance that participants received other notifications. The phone was then placed on the table, display down, next to the participant's dominant hand. Participants were told that they were not allowed to touch their phone during the experiment. During the last 20 trials of the last practice block preceding each Stroop version, the program sent three SMS messages, separated by seven seconds as not to be

mistaken for a call. Although participants could conclude their phone was set to vibrate, they could not know for certain that the notifications came from the experimenter.

In the control condition, participants were only instructed to set their phone to silent mode and put it in their pockets, as is customary at our lab. In addition, a notebook was placed next to their dominant hand. This way, the difference in performance between the two groups could not be attributed to just having an object next to one's dominant hand.

After the manipulation, participants proceeded to do the task. The task was coded in Python, version 2.7. The notifications in the notification condition were sent using the cloud-based messaging service Twilio (<https://www.twilio.com/>). After the task, participants reported several manipulation checks.

DEPENDENT MEASURE

Participants first performed a majority-congruent, followed by a no-congruent Stroop task. Each task had 144 randomly presented trials. In order to have an equal amount of trials across Stroop versions and trial types, we followed the procedure of Kane and Engle (2003); the program thus randomly determined 18 critical trials of each trial type before participants did the task. All analyses are conducted on the critical trials. Critical congruent trials paired each of the three color words (RED, BLUE, GREEN) with their respective ink colors six times. Critical incongruent trials paired each color word with the two conflicting colors three times. Critical neutral trials paired each word with the three colors two times. Neutral trials had the same length as color words, but did not share any letters (e.g., VHZT). Trials appeared in the center of the screen with a black background. Participants categorized the ink color by pressing one of three keys with their index, middle, or ring finger (counterbalanced). Before each task, participants completed three practice blocks (21 neutral trials each) to make them familiar with the task. Participants had to have 15 out of 21 practice trials for each block correct, otherwise the block would repeat. After the third block, participants were warned that they would not receive feedback anymore. Also, they were told they would now also encounter actual words rather than just the neutral trials.

To avoid inflated error rates at the beginning of the experimental blocks, participants first saw ten neutral trials (without feedback), followed by ten words. At the beginning of those 20 filler trials they received the three SMS if they were in the notification condition. This procedure was the same before each Stroop version. Each trial started with a fixation cross, displayed for 200ms. The colored word appeared immediately afterwards, and stayed on the screen for either 5s or until response. Intertrial interval was set to 500ms.

SAMPLING PROCEDURE AND PARTICIPANTS

Because previous work on the effect of smartphones on cognitive performance used different measures, we could not directly transfer previous effect sizes to calculate power. Instead, we employed a sequential Bayesian sampling design (Schönbrodt et al., 2017), which is efficient while maintaining superior error control compared to classical NHST

approaches. We set a minimum sample size of $n=20$ and a maximum of $n=50$ per condition (after exclusions), with a threshold of 3 and $1/3$, respectively, for the Bayes Factor (BF) for BF_{10} . We reached the threshold after the minimum sample size.

Participants ($N = 51$) were students from a university in the Netherlands, who received €5 or course credit. We observed interaction with cameras. Following our preregistration, we excluded six participants because they touched their phones during the experiment. No participants received any notifications other than the SMS sent by the experimenter. Furthermore, six participants were excluded because their phone did not vibrate, due to notification customizations the experimenter could not access. Nobody fulfilled our exclusion criterion of no more than 50% errors on neutral trials. The final sample consisted of 39 participants ($M_{age} = 22.00$, $SD_{age} = 2.97$, 31 females; $n_{notification} = 21$, $n_{control} = 18$).

STOPPING RULE

We calculated the difference between errors on neutral and incongruent trials (Stroop effect) separately for the majority-incongruent and no-congruent tasks. Then we calculated a difference score of these two Stroop effects (goal-neglect effect). To determine our stopping rule, we conducted a Bayesian independent t-test (after exclusions) on goal-neglect between the vigilance and control groups (JASP Team, 2017; the standard JASP prior was used, 0.707). We preregistered to stop collection at $BF_{10} = 3$ or $BF_{01} = 1/3$.

RESULTS

MANIPULATION CHECKS

As expected, everyone in the notification condition perceived the vibration of the SMS that they received. Participants in the notification condition also reported much higher ($M=59.57$, $SD=26.14$) distraction on a 100-point VAS compared to the control condition ($M=12.78$, $SD =22.02$; $t(36.99)=-6.07$, $p<.001$, $d=-1.94$). The same pattern emerged when comparing how distracting the notebook in the control condition was ($M=2.61$, $SD =2.78$) to the phone in the notification condition, $t(-20.52)=-9.92$, $p <.001$, $d=-3.07$). Further descriptive measures can be found on the OSF.

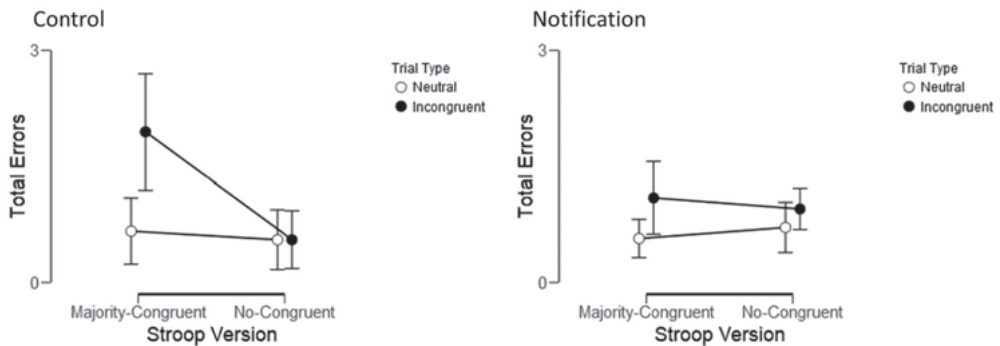


Figure 1. Line graph of the total errors per trial type and Stroop versions, separately for experimental condition. Lines represent 95% credible intervals.

PREREGISTERED ANALYSIS

Contrary to our expectation, participants in the notification condition had lower total errors on the goal-neglect measure ($M=.29$, $SD=1.42$) than participants in the control condition ($M=1.28$, $SD=1.64$); consequently, under the contrary assumption that we hypothesized, the data were about eight times more likely under the null hypothesis, $BF_{01}=8.34$, $95\%CI=[-.33, -.002]$, $d=.65$. Table 1 displays the average total errors for each condition by trial type and Stroop version. Apparently, the goal-neglect effect was only present in the notification condition compared to the control condition (see Figure 1), the opposite of what we hypothesized.

NONPREREGISTERED ANALYSES

The t-test we conducted for our stopping rule only gave us an overall test of the hypothesized three-way interaction, but was not informative with regard to whether we replicated the common Stroop or the goal-neglect effect. To further examine the exact contributions of each factor we conducted a 2 (vigilance: control, notification) \times 2 (Stroop version: majority-congruent, no-congruent) \times 2 (trial type: neutral, incongruent) Bayesian repeated-measures ANOVA, with the last two factors within, on total errors.

Inspecting the resulting table (Table 2), the second row demonstrates the Stroop effect with strong evidence (Lee & Wagenmakers, 2013) in favor of a difference between neutral and incongruent words compared to the null model ($BF_{10}=29$). In addition, the model with both main effects of trial type and Stroop version as well as their interaction ($BF_{10}=289$) outperformed the model without the interaction ($BF_{10}=289/63=4.59$). This constitutes moderate evidence for the goal-neglect effect, as the Stroop effect is more pronounced in the majority-congruent version compared to the no-congruent version. Thus, the goal-neglect model offers the best explanation for the data as it fits only slightly worse than the best fitting model ($BF_{10}=382/289=1.32$), yet is more parsimonious. Hence, the data

replicate the findings of Kane and Engle (2003). Last, the full model with the hypothesized three-way interaction fit better than the null model ($BF_{10}=157$), but did not offer substantial improvement over the goal-neglect model ($BF_{10}=157/289=0.54$). Thus, in accordance with the t-test we conducted for our stopping rule, there was only weak evidence for the inclusion of the three-way interaction across all models ($BF_{Inclusion}=2.50$).

Table 1. Mean and standard deviations for all groups of total errors on critical trials

Trial Type	Stroop Version	Condition	Mean	SD
Neutral	Majority-Congruent	Control	0.667	0.907
		Vigilance	0.571	0.598
	No-Congruent	Control	0.556	0.705
		Vigilance	0.714	0.784
Incongruent	Majority-Congruent	Control	1.944	1.798
		Vigilance	1.095	0.995
	No-Congruent	Control	0.556	0.616
		Vigilance	0.952	0.805

To better quantify the goal-neglect effect as proposed by Kane and Engle (2003), we conducted a Bayesian repeated-measures ANOVA without condition as a predictor. Indeed, the model with both main effects of trial type and Stroop version and their interaction was the best fitting model ($BF_{10}=287$) and there was moderate evidence for the interaction effect ($BF_{10}=287/65=4.42$) with a considerable effect size ($\eta^2_p=.18$, $\omega^2=.06$).

To further investigate the weak evidence for a three-way interaction and gain more insight into the nature of the goal-neglect effect, we conducted separate Bayesian ANOVAs for each condition. The control condition (Figure 1) clearly showed the expected pattern of goal-neglect, as the model with the interaction displayed superior fit compared to a model with only the main effects ($BF_{10}=215/35=6.14$). In contrast, a separate ANOVA for only the vigilance condition demonstrated weak support for any of the effects, with the best fitting model only containing trial type as predictor (i.e., the Stroop effect, $BF_{10}=2.87$); instead, there was weak evidence for the lack of an interaction effect when comparing the full model to a model with two main effects ($BF_{01}=3.65/1.57=2.32$), which by itself displays weak fit.

Table 2. Bayesian mixed ANOVA

Model Comparison					
Models	P(M)	P(M data)	BF_M	BF₁₀	error %
Null model (incl. subject)	0.053	7.781e -4	0.014	1.000	
Trial Type	0.053	0.023	0.418	29.190	1.134
Stroop Version	0.053	0.001	0.026	1.858	1.070
Trial Type + Stroop Version	0.053	0.050	0.942	63.932	1.754
Trial Type + Stroop Version + Trial Type * Stroop Version	0.053	0.225	5.236	289.628	2.564
condition	0.053	2.006e -4	0.004	0.258	2.505
Trial Type + condition	0.053	0.006	0.104	7.373	1.546
Stroop Version + condition	0.053	3.579e -4	0.006	0.460	1.250
Trial Type + Stroop Version + condition	0.053	0.014	0.251	17.670	4.571
Trial Type + Stroop Version + Trial Type * Stroop Version + condition	0.053	0.056	1.075	72.430	3.182
Trial Type + condition + Trial Type * condition	0.053	0.002	0.034	2.439	4.495
Trial Type + Stroop Version + condition + Trial Type * condition	0.053	0.004	0.076	5.424	5.054
Trial Type + Stroop Version + Trial Type * Stroop Version + condition + Trial Type * condition	0.053	0.020	0.365	25.534	5.704
Stroop Version + condition + Stroop Version * condition	0.053	0.001	0.024	1.738	4.312
Trial Type + Stroop Version + condition + Stroop Version * condition	0.053	0.059	1.130	75.886	2.721
Trial Type + Stroop Version + Trial Type * Stroop Version + condition + Stroop Version * condition	0.053	0.298	7.634	382.735	3.943
Trial Type + Stroop Version + condition + Trial Type * condition + Stroop Version * condition	0.053	0.019	0.346	24.213	2.858
Trial Type + Stroop Version + Trial Type * Stroop Version + condition + Trial Type * condition + Stroop Version * condition	0.053	0.098	1.965	126.515	3.867
Trial Type + Stroop Version + Trial Type * Stroop Version + condition + Trial Type * condition + Stroop Version * condition + Trial Type * Stroop Version * condition	0.053	0.122	2.503	156.919	6.601

Note. All models include subject.

DISCUSSION

This experiment demonstrated the common Stroop effect, such that incongruent words produced more errors compared to neutral trials. Furthermore, it is encouraging that we replicated the findings of Kane and Engle (2003) with a preregistered study. The Stroop versions produced a strong interaction effect, with an effect size in the same range as their work. It indeed appears that a majority-congruent Stroop task leads to neglecting the goal of identifying the ink compared to a no-congruent Stroop task. Replicating both the Stroop and the goal-neglect effect lends credibility to our measurement.

With regard to our manipulation, the notifications were perceived as very distracting in comparison to the control group. This finding replicates our previous study (Johannes et al., in press). In addition, although we did not employ a direct vigilance measure here (because the study reported here was conducted earlier than Johannes et al.), this exact manipulation has now been shown to influence vigilance (Johannes et al., in press). In addition, the Bayesian sequential sampling design we employed in general needs smaller samples compared to a classic NHST approach to reach a conclusion about an effect. Thus, it is sensitive to detecting whether the data support or do not support an effect, given a certain prior, even with small samples such as ours, while performing well in terms of long-term error control (Schönbrodt et al., 2017). Taken together, those results imply that our experimental manipulation worked and our task measured what it was intended to measure, while the analysis was sensitive to detect an effect.

It is important to note that Bayesian sequential sampling is not yet commonly used and there are no clear conventions on setting the BF cut-offs. We chose 3, which reflects anecdotal evidence and can thus be considered a relatively low evidential threshold (Lee & Wagenmakers, 2013). However, BFs have the advantage that readers can decide for themselves whether they find cut-off and evidential value convincing. Therefore, we recommend future research choose cut-offs based on the literature and what they themselves consider convincing evidence.

Contrary to our hypothesis, there was moderate evidence that vigilance did not lead to stronger goal-neglect. If anything, our data suggest the opposite. In the control condition, we observed the typical goal-neglect effect. For the notification condition, we expected that this pattern would be more pronounced. However, we found no evidence for the goal-neglect effect; instead, our analysis suggests the lack of an effect. There was also no evidence for the general Stroop effect here. Thus, we can conclude that vigilance did not interfere with goal-neglect; if anything, it might have even lead to improved performance, reducing the Stroop effect and eliminating the goal-neglect effect.

Interestingly, the lack of effect can be explained by the same account that also explains a lack of effect on response inhibition (Johannes et al., in press). It is possible that a mindset of avoidance was induced because participants were not allowed to touch their phones. This mindset might have raised arousal and recruited additional resources which has shown to improve performance on a Stroop task (Koch et al., 2008). Thus, smartphone

vigilance might have had both a positive and a negative effect on working memory. On the one hand, the state might have been taxing to executive functioning. On the other hand, the increased arousal might have recruited cognitive resources that helped to keep the goal in working memory, ultimately evening out the negative effect of smartphone vigilance. Yet, it is not clear if and how this avoidance-induced alertness interacts with or is theoretically similar to the alertness of smartphone vigilance.

Taken together, there is increasing evidence that smartphone vigilance is experienced as distracting, but does not impair cognitive control. Consequently, smartphone vigilance impairing working memory does not provide a viable explanation for previous work finding negative effects of smartphone notifications on cognitive performance and attention (Stothart et al., 2015; Thornton et al., 2014; Ward et al., 2017). Furthermore, it appears necessary to distinguish between constant interruptions (in the form of checking behavior) and the mindset of vigilance. Whereas actual behavior, for example multitasking, has shown to impair executive functions (Baumgartner et al., 2017), situations where smartphones are considered temptations might not be as detrimental as previously assumed by both scholars and users. Our findings call for more research differentiating smartphone interaction and smartphone vigilance.

Social Smartphone Apps Do Not Capture Attention Despite Their Perceived High Reward Value

This chapter has been published as Johannes, N., Dora, J., & Rusz, D. (2019). Social smartphone apps do not capture attention despite their perceived high reward value. *Collabra: Psychology*, 5(1), 14. <https://doi.org/10.1525/collabra.207>.



ABSTRACT

Smartphones have been shown to distract people from their main tasks (e.g., studying, working), but the psychological mechanisms underlying these distractions are not clear yet. In a preregistered experiment (<https://osf.io/g8kbu/>), we tested whether the distracting nature of smartphones stems from their high associated (social) reward value. Participants (N = 117) performed a visual search task while they were distracted by (a) high social reward apps (e.g., Facebook app icon + notification sign), (b) low social reward apps (e.g., Facebook app icon), and (c) no social reward apps (e.g., Weather app icon). We expected that high social reward app icons would slow down search, especially when people were deprived of their smartphones. Surprisingly, high social reward (vs. low or no social reward) apps did not impair visual search performance, yet in a survey (N = 158) participants indicated to perceive these icons as more rewarding. Our results demonstrate that even if people perceive social smartphone apps as more rewarding than nonsocial apps, this may not manifest in behavior.

Smartphones are thought to be pervasive sources of distractions, defined as performance decrements after the onset of task-irrelevant stimuli (Rusz, Bijleveld, & Kompier, 2018). Indeed, increasing experimental evidence shows that smartphones impair cognitive performance (Chein et al., 2017). For instance, hearing a phone ring (Shelton et al., 2009), receiving notifications (Stothart et al., 2015), or even the mere presence of a smartphone (Thornton et al., 2014; Ward et al., 2017) had a negative effect on sustaining attention on a main task (but see also Johannes et al., in press). In line with such an impairment in maintaining attention, Kushlev, Proulx, and Dunn (2016) found that people report more difficulties to concentrate on their tasks when they enable (vs. disable) notifications. Taken together, there is growing experimental evidence that smartphones appear to harm productivity. However, the underlying psychological mechanism of these performance decrements remains unknown. Understanding this mechanism is crucial, as it can advance theory on the effects of smartphones on performance and inform policy makers on how to deal with smartphone use, for instance in school or work contexts.

Previously, smartphone distractions have predominantly been explained as a stimulus-driven mechanism. From this perspective, impairments in performance happen because people are distracted by an external source (e.g., notifications, ringing phone). However, such a perspective does not explain why a smartphone notification should have a stronger effect than any other external stimulus (e.g., a loud tone). Instead, people are not only influenced by external cues, but also driven by current motivational states (Botvinick & Braver, 2015). Therefore, beyond external sources, smartphone distractions can be explained by a motivational drive to seek social rewards.

In line with this idea, it is plausible that smartphones distract people from their tasks because they carry social reward to the user and the user is motivated to attain that reward despite disengaging from another task (Oulasvirta et al., 2012). According to Bayer, Campbell, and Ling (2015), because people have an innate need for social contact and belonging (Baumeister & Leary, 1995; Deci & Ryan, 2000), they use the predominantly social features of smartphones such as WhatsApp or Facebook. Through repeatedly meeting their social needs on those apps, users form an association between social reward and their smartphones. Thus, users are first motivated to attain social rewards through their smartphones. Once this connection is established, smartphone cues, such as receiving a notification, may automatically attract attention and trigger checking habits. In sum, Bayer and colleagues (2015) assume that the distracting potential of smartphones is due to their rewarding nature.

Although this account appears plausible, there are no direct tests of a smartphone cue-reward association. As of now, most research relies on indirect tests. For instance, there is evidence that smartphone symbols are associated with positive affect (van Koningsbruggen et al., 2017) and can prime relationship-related concepts (Kardos et al., 2018). Additionally, there is ample cross-sectional evidence demonstrating that users themselves report that they obtain social gratification from social apps (Ishii, Rife, & Kagawa, 2017; Jung & Sundar, 2018; Karapanos et al., 2016). Thus, even though several studies have addressed the idea

that smartphones are associated with high social rewards, there is no direct empirical test of this mechanism.

On a fundamental level, value-driven attention (for a review see Anderson, 2016b) provides a well-established cognitive framework that can explain reward associations, including those with one's smartphone. As people, by nature, are reward-seeking organisms (Braver et al., 2014), attention prioritizes information that signals reward (Chelazzi et al., 2013). Recent work shows that this prioritization process operates even when information is entirely task-irrelevant, which leads to disengagement from the task at hand (Anderson et al., 2011a; Rusz et al., 2018). In a series of studies (Anderson et al., 2011a; Anderson, Laurent, & Yantis, 2011b; Le Pelley et al., 2016; Theeuwes & Belopolsky, 2012), participants first learned to associate an arbitrary stimulus feature (e.g., color) with high or low monetary rewards. Later, they engaged in a visual search task where these colored stimuli appeared as nontargets that needed to be ignored. Results show that distractors associated with high (vs. low) monetary rewards significantly slowed down visual search. This means that reward-associated distractors gain high attentional priority (i.e., become more salient) and therefore capture visual attention (Hickey, Chelazzi, & Theeuwes, 2010). This mechanism of learning to associate rewards with certain stimuli could explain how reward associations take place in smartphone settings.

VALUE-DRIVEN ATTENTION AND SMARTPHONE APP ICONS

Applying a value-driven attentional mechanism to a smartphone scenario, it is plausible that certain smartphone features (e.g., app icons) have been associated with social rewards through repeated use. Consequently, these features gain attentional priority and therefore attract attention and eventually harm visual search performance. As the major part of social interaction on smartphones happens via apps, we assume that app icons carry social reward to the user. For instance, social apps (e.g., Facebook, WhatsApp), particularly with a notification sign, should be associated with high social reward, as notifications usually convey social validation, such as friends liking a picture or friend requests (Reich et al., 2018). Conversely, nonsocial apps (e.g., Weather, Calculator) should not carry social rewards as they are not used for social purposes. So, analogous to the value-driven attention account, we expect that social app icons should similarly attract attention and therefore slow down visual search. Therefore, we predict that low social reward distractors (social app icons) and high social reward distractors (social app icons with a notification) result in slower reaction times compared to no reward distractors (neutral app icons; H_{1a-b}), and that high social reward distractors result in slower reaction times than low social reward distractors (H_{1c}).

In addition, it is well-established that deprivation of rewarding experiences strengthens the motivation to obtain these experiences (Seibt et al., 2007). For example, depriving participants of food led to a higher reinforcing value of the food compared to not hungry participants (Epstein, Truesdale, Wojcik, Paluch, & Raynor, 2003). Similarly, it is common practice to assess the true value participants assign to food after a fasting period (Z. Chen

et al., 2016). In the case of smartphones, if social apps truly are rewarding, the appeal of social apps, similar to food, should be stronger for those who have been deprived of using these apps. Evidence for such a position comes from studies showing that phone separation is associated with strong emotional reactions (Hoffner & Lee, 2015), leads to anxiety (Cheever, Rosen, Carrier, & Chavez, 2014), impairs cognitive control (Hartanto & Yang, 2016), and results in physiological stress reactions (Clayton, Leshner, & Almond, 2015). Consequently, the reward value associated with app icons should be particularly high, and hence distracting, when participants are motivated to use these apps. We thus hypothesize that all main effects of distractor are stronger for users who have previously been deprived of their phones compared to a control group (H_2).

This set-up enables us to exclude alternative explanations: If we indeed find the expected pattern for distractor, (a) low and high social reward distractors might capture attention merely because the social apps are more familiar to participants, given that they are used more; (b) high social reward distractors might capture attention more than low social reward distractors because of the red color of the notification sign. Therefore, only if the effect is amplified in the deprivation condition can we conclude that apps indeed carry reward for users, above and beyond the possible effects of familiarity and color.

To test our hypotheses, we adapted the visual search task introduced in Anderson et al. (2011b). We chose this paradigm for two reasons. First, it is a well-established method to assess the effect of reward-associated distractors on attention (for reviews see Anderson, 2016b; Failing & Theeuwes, 2017; Le Pelley et al., 2016). Second, the visual search task represents a good approximation of smartphone distractions in real life scenarios. For instance, consider a student who has to write a paper, but the Facebook notification sign repeatedly captures their attention.

We deviated from the original paradigm in two major aspects. First, we omitted the reward learning phase from the current study because we assumed that people learned to associate social rewards with smartphone app icons through repeated exposure in everyday life. Therefore, we only used the testing phase of the original paradigm. Second, in order to increase ecological validity, we used smartphone app icons as distractors. By using real-life icons, we followed recent studies which show that more complex visual information, such as pictures of people or scenery, can also be associated with rewards (Failing & Theeuwes, 2017; Hickey, Kaiser, & Peelen, 2015).

Thus, in the current study, participants were instructed to find the target while they were distracted by app icons that were associated with high social rewards, low social rewards, or no social rewards. In the original paradigm, the rewarding nature of stimuli is reflected in impaired visual search. Consequently, the visual search task paradigm provides us with a test of the proposed smartphone-reward association: If social smartphone cues are indeed more rewarding than neutral smartphone cues, they, like other rewarding stimuli, should impair visual search. In other words, impaired visual search performance serves as an indicator of the reward associated with smartphone cues.

STUDY 1

METHOD

Preregistration and Data Availability. We preregistered hypotheses, sample size, inclusion and exclusion criteria, and statistical analyses. Our preregistration, experimental materials, data, and analysis are available on the Open Science Framework (<https://osf.io/g8kbu/>).

PARTICIPANTS

As power calculations are not entirely straight-forward for linear mixed-effects models (Scherbaum & Ferrer, 2008), we preregistered a rather conservative sample size. Therefore, we recruited 120 students from a Dutch university. We had four inclusion criteria: First, participants needed to have normal or corrected to normal vision. Second, students needed to own an iPhone. This ensured the icons we used as distractors would be identical to those that participants use on their iPhones every day. Icons are standardized across iOS compared to Android, where icons often differ between devices due to the open source nature of the Android OS. Third, as people under 25 report the highest smartphone use (CBS, 2018; Pew Research Center, 2017), our participants had to be younger than 25 years. Fourth, participants had to have the five distractor apps Facebook, Facebook Messenger, Instagram, Snapchat, and WhatsApp installed on their iPhone and they had to be frequent users of these apps for at least two years. These criteria were meant to ensure that reward learning had taken place, that is, stimulus features had been paired with the delivery of (social) rewards (Le Pelley et al., 2016): Using these five social apps frequently plausibly has led to an established association of social rewards with visual features of these apps.

Following our preregistered exclusion criteria, we excluded three participants as they did not reach 70% accuracy on the task. Thus, the final sample consisted of 117 students (59 in the control and 58 in the deprivation condition; 106 females, $M_{age} = 20.85$, $SD_{age} = 1.88$). Participants were compensated with monetary rewards in the form of a gift voucher (€5 or €10) or course credits. The study had IRB approval and all participants gave informed consent.

DESIGN

We employed a mixed design with deprivation as a between-subject independent variable (2 levels: deprivation group vs. control group), app distractor icon as a within-subject independent variable (3 levels: high social reward vs. low social reward vs. no social reward) and response time (RT) as dependent variable.

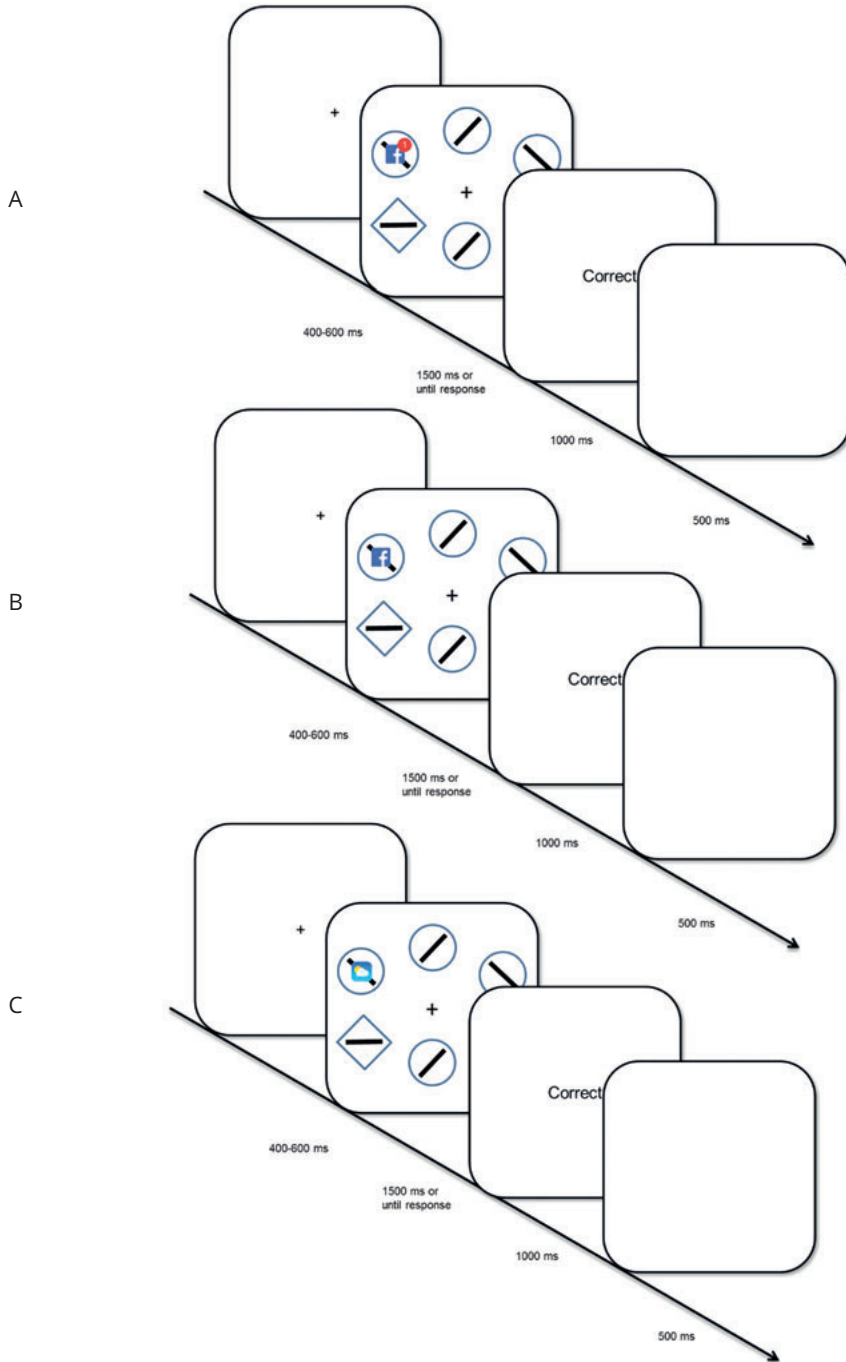


Figure 1. Trials in the experiment. Examples of (A) high social reward distractor trial, (B) low social reward distractor trial, and (C) no social reward distractor trial.

PROCEDURE

We randomly assigned participants to either the deprivation or the control condition. In the *deprivation condition*, we asked participants to come to the lab *one hour* before the experiment to hand in their iPhone, which we locked away in a drawer. Then, we told participants that they were free to go about their day within the next hour, but asked them not to engage in any social media activity until the experiment started. After one hour, they came back and performed the task (see below). After the task they received their phone. In the *control condition*, participants came to the lab at their assigned time slot and directly performed the task.

Before starting the task, participants reported demographics (age and gender). In order to assess whether our deprivation manipulation indeed led to an increased motivation to use their smartphones, participants then answered a short manipulation check on a 1 (*not at all*) to 100 (*extremely*) visual analogue scale ("Right now, to what extent do you feel an urge to check your phone?"). Then, they performed the visual search task. Finally, after finishing the task, participants reported a second manipulation check, namely whether they had seen 20 apps (ten of which were used in the experiment) during the course of the visual search task. With this question, we tested whether participants actually processed the distractor app icons throughout the visual search task.

VISUAL SEARCH TASK

We designed a visual search task based on Anderson et al. (2011b). Participants were seated about 50 cm from a monitor with a resolution of 1920x1080 pixels. On each trial, participants first saw a fixation cross with a visual angle of 0.5°, then six shapes organized in an imaginary circle with a visual angle of 10°; each shape had a visual angle of 3.45°; last, participants were presented with a performance feedback display (see Figure 1). Among these six shapes, there was always one unique shape, which was defined as the target (i.e., a circle among diamonds or a diamond among circles). Each nontarget shape contained a black line tilted by 45°. The target shape always contained either a horizontal or vertical black line. On all trials, there was a distractor app icon embedded (1.73° visual angle) in one of nontarget shapes, on top of the tilted lines.

These distractor app icons represented three levels of social rewards (high, low, and no social rewards, see Figure 2). On the high social reward distractor trials (see Figure 2A), there was a *social app icon with a notification sign* (Facebook, Facebook Messenger, Instagram, WhatsApp, and Snapchat) within one of the nontarget shapes. We chose these apps because they are the most commonly used social apps. The red notification was identical to the one used on iOS. On the low social reward distractor trials (Figure 2B), there was a *social app icon* (i.e., same icons without the notification sign) within one of the nontarget shapes. As stated above, these apps are mainly used for social purposes – so we assumed they represent social reward to people, but less than these same apps with the certainty of a notification sign. Finally, on the no social reward distractor trials (Figure 2C), there was a *neutral app icon* (Weather, Settings, Notes, Clock, and Calculator) within the

nontarget shapes. We chose these specific icons as they are pre-installed on every iPhone, so iPhone users most likely encounter them often enough; yet, they are never used for social purposes, so we assumed that participants could not have possibly associated social rewards with any of the neutral app icons. The target shape never included any distractors (i.e., icons). Target and distractor location were randomly determined; distractor app icon and the unique shape were counterbalanced.

Participants were instructed to search for the target, which was always defined as the unique shape in the search display, and report whether the line within the target shape was horizontal or vertical, by pressing the “z” and “m” keys (counterbalanced). The experiment consisted of 480 trials: 120 trials (25%) contained a high social reward distractor, 120 trials (25%) contained a low social reward distractor, and 240 trials (50%) contained a no social reward distractor. Before the task, participants did 24 practice trials. After each 96 experimental trials, participants could take a short break. The task took ~ 35 minutes to finish.



Figure 2. Stimuli used in the experiment. (A) social app icons with a notification sign represent high social rewards. (B) social app icons represent low social rewards. (C) neutral app icons represent no social rewards.

DATA ANALYSIS

We conducted all of our analyses in R (version 3.5.0, R Core Team, 2018). In line with our preregistration, we tested our hypotheses using a linear mixed-effects modeling approach using the *lmer* function (lme4 package; version 1.1.17; Bates et al., 2015). We aimed for a ‘maximal’ random effects structure as advocated by Barr, Levy, Scheepers, and Tily (2013) to avoid inflated Type-1 errors. Accordingly, our model predicting response time

included two random intercepts; a per-participant random intercept to account for the repeated-measures nature of the data and a per-app icon random intercept to account for any additional variance in response time caused by the specific app icons included in our study. We modeled the within-subject predictor distractor as fixed effect and as random slope varying across participants. We modeled the between-subject predictor condition as fixed effect and as random slope varying across app icons.

To determine p-values, we preregistered to compute Type III bootstrapped Likelihood Ratio Tests using the *mixed* function (afex package; version 0.20-2; Singmann et al., 2018). However, this analysis led to several convergence warnings that persevered after the recommended troubleshooting steps. Thus, we followed recent recommendations by Luke (2017). Based on simulations, he compared several approaches to evaluating significance in mixed-effects models, and concluded that F-tests with Satterthwaite approximation for degrees of freedom are the most appropriate to control Type 1 error rates. Thus, we opted for this approach instead (also using the *mixed* function), which resulted in no convergence warnings.

RESULTS

MANIPULATION CHECKS

Directly before starting the visual search task, participants in the deprivation condition reported a higher urge to check their smartphone ($M = 51.81$, $SD = 21.20$) than participants in the control condition ($M = 32.28$, $SD = 27.15$), $t(111.44) = -4.39$, $p < .001$, $d = .80$. At the end of the experiment, participants correctly classified whether or not they had seen 20 different app icons (ten of which we used as distractors) with an accuracy of 84%, indicating that they did process the distractors during the search task.

PREREGISTERED ANALYSES

In line with our preregistration, we excluded any trial on which (a) the RT was below 300ms ($< 0.01\%$) and (b) the RT was ± 3 SDs from the participant's mean (0.01%). For the analysis, we also excluded all inaccurate trials. Participants were accurate on 92% of the experimental trials. Across all remaining experimental trials from all participants ($N = 51083$) mean response time was 676.46ms ($SD = 81.17$).

The main effect of distractor was not significant, $F(2, 13.49) = 0.90$, $p = .428$. Participants' response time did not significantly differ between high social reward distractors ($M = 678.82$, $SD = 83.11$), low social reward distractors ($M = 676.52$, $SD = 82.06$), or no social reward distractors ($M = 675.26$, $SD = 81.28$). To our surprise, the main effect of condition was significant, $F(1, 114.99) = 4.00$, $p = .048$. Overall, participants in the deprivation condition ($M = 661.61$, $SD = 76.17$) responded faster than participants in the control condition ($M = 691.06$, $SD = 83.88$), irrespective of the type of distractor presented on any given trial. Last, the interaction effect of distractor and condition was not significant, $F(2, 348.63) = 2.59$, $p = .076$. To investigate whether there was indeed no interaction effect and to better understand our data, we tested the main effect of distractor in both conditions separately.

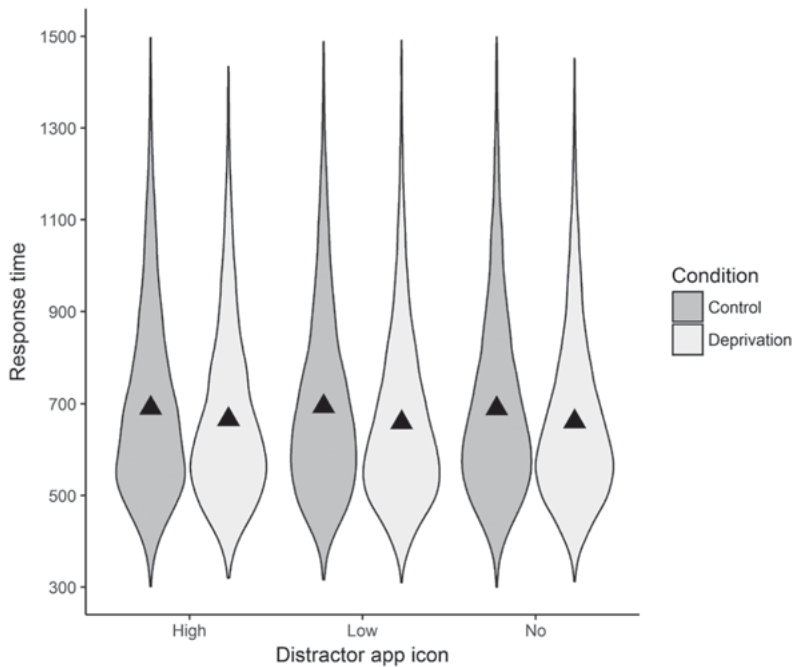


Figure 3. Violin plots of response times per distractor and condition. Triangles represent mean response times (in ms).

The main effect of distractor was neither significant in the control condition, $F(2, 16.54) = 1.18, p = .33$, nor in the deprivation condition, $F(2, 16.18) = 2.56, p = .11$. Taken together, the effect of distractor did not significantly differ between the deprivation condition and the control condition. A visualization of the raw data associated with our analysis can be found in Figure 3.

BAYESIAN FOLLOW-UP ANALYSES

A major limitation of our frequentist model is that it cannot quantify evidence for the null hypothesis. Therefore, to investigate to what extent our data support the lack of an effect, we conducted a Bayesian repeated-measures ANOVA with the *anovaBF* command (BayesFactor package; version 0.9.12-2; Morey & Rouder, 2015). The model employed the default Cauchy distribution for the prior. The Bayes Factors associated with our predictors can be found in Table 1. Comparing a model with the main effect of condition to the null model yielded inconclusive evidence, as the data were 1.63 times more likely under the null model without the effect of condition ($BF_{10} = 0.61$). On the one hand, the Bayesian ANOVA does not allow an analysis as fine-grained as the frequentist mixed model, as it does not include a per-icon random intercept and random slope of condition. On the other hand, p-values close to the cut-off of $\alpha = .05$ often do not represent much evidential value (Benjamin et al., 2018), which is further illustrated by the Bayes Factor we obtained. The

Bayesian analysis of the main effect thus shows that we should interpret the significant main effect of deprivation with caution.

In addition, supporting the nonsignificant effect of distractor, there was strong evidence that the data were much more likely under a null model compared to a model with the effect of distractor ($BF_{01} = 755$). The same holds for the interaction effect, which was not supported compared to a model with the two main effects ($BF_{01} = 155$).

Table 1. Results of Bayesian follow-up analysis

Effect	BF
Condition	0.612712
Distractor	0.001325
Condition + Distractor	0.000777
Condition + Distractor + Interaction	0.000005

EXPLORATORY ANALYSES.

In order to follow up on the unexpected main effect of condition, we investigated whether there was a speed-accuracy tradeoff. A maximal generalized mixed-model with accuracy as the dependent variable did not show a significant effect of condition ($\chi^2(1) = 0.0$, $p = .99$). Supporting the lack of an effect, a Bayesian contingency table showed strong support for the lack of a difference between the conditions ($BF_{01} = 167$). We conclude that there was no speed-accuracy tradeoff, and that participants in the deprivation condition indeed performed better (faster while equally accurate).

DISCUSSION

Contrary to our expectations, high social reward apps did not slow down visual search compared to low or no social reward apps, neither in the smartphone deprived, nor in the control condition. Based on prior work we assumed that different apps would have different levels of reward associated with them (e.g., Bayer et al., 2015; van Koningsbruggen et al., 2017). However, one possible explanation for this null effect is that social apps were not perceived as more rewarding than neutral apps. In fact, unlike in the original study series on value-driven attention, we did not directly manipulate stimulus-reward associations. In the original task, participants go through an extensive reward training, in which arbitrary stimuli, such as color, become associated with the delivery of monetary rewards. Consequently, these reward-associated stimuli slow down visual search; that is, impairment of visual search is an indicator of attentional capture by the reward of the stimuli. However, in our application of this paradigm we did not manipulate reward, but assumed the reward value of apps had been established in real life, through repeated use prior to the experiment. The lack of an effect on visual search speed might then either

reflect that the stimuli are not rewarding, or that they are rewarding, but not rewarding enough to cause differences in attentional capture. Due to the design of Study 1, we cannot be certain that participants indeed perceived social apps as more rewarding than nonsocial apps. Therefore, we need to establish whether our reward manipulation was effective after all to rule out the alternative explanation that the app categories were not different in their associated reward.

To address this possible alternative explanation for the null effect of reward-associated distractors, we conducted a survey where participants rated all 15 apps we used during the experiment on how rewarding they found them. We expected that, if the three levels were truly to manipulate social reward, we should at least be able to detect a difference on how people themselves perceive these different apps. Accordingly, we hypothesized that high social reward apps would be rated higher than both low social reward apps and no social reward apps. In addition, we expected low social reward apps to receive higher ratings than no social reward apps.

STUDY 2

METHOD

Preregistration and Data Availability. We preregistered hypotheses, sample size, inclusion and exclusion criteria, and statistical analyses. The preregistration, experimental materials, data, and analysis are available on the Open Science Framework project of this article (<https://osf.io/g8kbu/>).

PARTICIPANTS

Because we expected an experimental manipulation to induce at least a medium-sized effect ($\eta_p^2 = .05$) on a manipulation check, we aimed to obtain 95% power to detect an effect of at least that size at $\alpha = .05$ for the main effect in a repeated-measures ANOVA. Thus, we preregistered to recruit 160 (150 needed for 95% power plus ten to account for exclusions) valid responses on the online platform Prolific. We counted those submissions as valid that passed an attention check (see below), as Prolific lets researchers resample participants if a participant fails an attention check.

We aimed to obtain a sample as similar as possible to our sample in Study 1. Overall, 252 participants from the UK between the ages of 18 and 25 opened the survey. All participants were screened and had to currently own an iPhone and have used an iPhone for at least the past two years. Furthermore, participants had to have the five social apps from Study 1 installed and had to have used them for at least the past two years. In addition to these inclusion criteria, we preregistered several exclusion criteria. First, 52 participants were excluded because they did not finish the survey. Second, of the remaining 200, 40 did not pass an attention check (see Procedure). Third, we excluded two participants who

indicated that they were older than prescreened by Prolific. No participant fulfilled our fourth exclusion criterion of having a variance of zero across all rated apps, or our fifth exclusion criterion of spending less than 30 total seconds on the 15 apps to rate ($M_{\text{seconds}} = 72$, $SD_{\text{seconds}} = 31$). Thus, the final sample consisted of 158 participants ($M_{\text{age}} = 21.56$, $SD_{\text{age}} = 2.40$) of which 110 were female (70%).

PROCEDURE

Participants were informed that the aim of the study was to find out how people experience different apps. In particular, participants were informed that they were to rate different apps on how rewarding they find them. To make clear what we meant with rewarding, we provided several clarifications (e.g., feeling happy when using the app, feeling a strong need to use it, liking the app). To avoid participants overthinking their responses, we instructed them to respond promptly, based on their immediate thoughts about each app. To avoid confusion about the difference between a high social reward app (i.e., a social app with a notification sign) and a low social reward app (i.e., the same social app without a notification sign), we instructed participants that the apps would sometimes have a notification sign and that they should treat the app as if they saw it in that form on their own phone. Because understanding the task instructions was crucial to accurately rate the apps, we implemented two measures to ensure participants properly read the

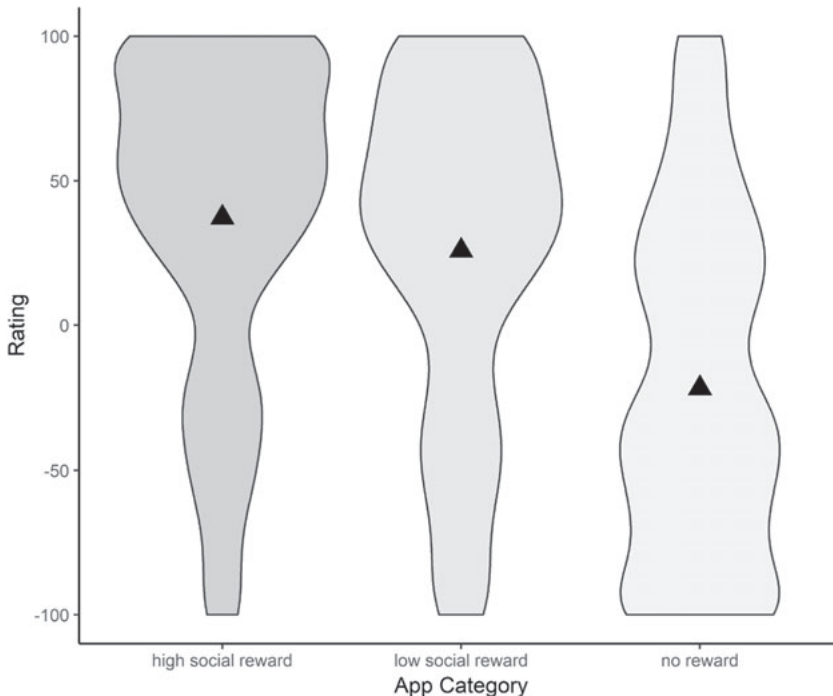


Figure 4. Distribution of how rewarding participants rated the three categories of apps. Triangles represent mean ratings.

instructions. First, going to the next page was only possible after 20 seconds. Second, at the end of the task description, we instructed participants to select “No” to proceed to the task as an attention check.

Participants then proceeded to rate all 15 stimuli used in Study 1 on the question “How rewarding do you find this app?” on a visual analogue scale ranging from -100 (*not at all*) to 100 (*very much*). Presentation order of the apps was randomized. The entire survey, on average, took about three minutes ($M_{\text{seconds}} = 185$, $SD_{\text{seconds}} = 71$) and participants received £0.50. The study had IRB approval and all participants gave informed consent.

RESULTS

We conducted a repeated-measures ANOVA with app category (within: high social reward vs. low social reward vs. no social reward) as predictor and ratings of how rewarding participants found those apps as outcome. As the assumption of sphericity was violated ($W = .30$, $p < .001$), we report the F -statistic with Greenhouse-Geisser correction. The main effect of category was significant and large, $F(1.18, 184.75) = 150.77$, $p < .001$, $\eta^2 = .32$. The strength of evidence for an effect of category was further supported by a Bayesian repeated-measures ANOVA with the standard Cauchy prior ($BF_{10} = 1.63e+107$). To test our predicted contrasts we conducted three post-hoc two-tailed paired t -tests without correction for multiple testing, as correction for multiple testing is not necessary for designs with only one factor with three levels. We present paired Bayesian t -tests alongside the frequentist results.

In line with our predictions, high social reward apps ($M = 36.99$, $SD = 33.18$) received significantly higher ratings than low social rewards apps ($M = 25.46$, $SD = 34.31$), $t(157) = 7.61$, $p < .001$, $BF_{10} = 3.06e+09$, $d_z = 0.61$, and significantly higher ratings than no social reward apps ($M = -22.00$, $SD = 43.48$), $t(157) = 13.20$, $p < .001$, $BF_{10} = 1.15e+24$, $d_z = 1.05$. In addition, low social reward apps received significantly higher ratings than no social reward apps, $t(157) = 11.64$, $p < .001$, $BF_{10} = 7.36e+19$, $d_z = 0.93$. The residuals within each condition were roughly normally distributed and the results were robust to removal of outliers. A visualization of the raw data associated with our analysis can be found in Figure 4.

GENERAL DISCUSSION

The goal of the current study was to test whether smartphone distractions stem from the high social rewards associated with smartphone apps. Participants engaged in a visual search task while they were distracted by smartphone app icons. Although we show that participants *perceive* social apps as more rewarding than neutral apps, that perceived reward did not impair performance in a visual search task. Also, depriving participants of their smartphone did not amplify such an effect. However, surprisingly, participants who were deprived of their smartphones performed better. In short, these results suggest that even if people perceive social apps as more rewarding than nonsocial apps, being exposed

to these apps as distractors does not influence performance on a visual search task. However, there are several alternative explanations, both theoretical and methodological, for our findings.

One possible alternative explanation for the lack of an effect of the three app groups is that participants did not perceive social apps as more rewarding than neutral apps. Instead of manipulating reward, as is common with the visual search paradigm, we assumed that users repeatedly obtain social validation and gratification from social apps (Karapanos et al., 2016; Reich et al., 2018), such that they would learn to associate social reward value with these apps. Therefore, we expected that social app icons, particularly those with a notification sign, gained their reward value in everyday life and should be as powerful as a controlled reward training phase in the lab. To provide evidence for this line of reasoning, Study 2 showed that people themselves report social apps to be more rewarding than neutral apps, especially if social apps have a notification sign. Importantly, the effect we obtained was large. As a consequence, we can be more confident that the lack of an effect is not due to a failed manipulation of reward.

That being said, there are several caveats to this objection which do not allow to draw a clear conclusion from our behavioral data regarding the reward value of apps. First, Study 1 and Study 2 were run on different samples, albeit matched on demographics. Technically, insights from the ratings in Study 2 might not apply to participants in Study 1. In addition, although we show a difference in how the different app sets (i.e., high, low, and no social reward) are perceived, these ratings are relative to each other. We cannot be certain a social app with a notification sign does truly feel rewarding -- or just more rewarding compared to a neutral app, which might not feel rewarding at all. As such, the relative difference in perceived reward value might not manifest itself on a behavioral level because, in absolute terms, the reward associated with apps is not large enough to attract attention (Potter, 2011).

Furthermore, we did not have a no-app control condition. Such a control condition would be informative by testing whether all apps, regardless of their perceived value, slow down visual search. Similarly, implementing a control condition with an arbitrary symbol (e.g., a symbol similar in shape to app icons) as distractor could provide a test whether app icons attract attention above and beyond *any* other distractor. This view aligns with the lack of an interaction between app icons and the deprivation manipulation. We predicted that social apps would be particularly distracting if users had been deprived to access them (Epstein et al., 2003; Seibt et al., 2007). Yet our data show that deprivation did not affect whether participants were more or less distracted by different apps. The lack of an interaction provides additional evidence for the explanation that perceived reward did not manifest itself on a behavioral level. Future research could consider using a no-reward control condition, an arbitrary symbol control condition, or even contrast apps with the low and high monetary reward condition used in the original paradigm (Anderson et al., 2011b) to test such a proposition.

Our findings echo work demonstrating that people's perception about smartphones do not necessarily translate to behavior. For instance, Johannes et al. (in press) found that receiving a notification during a cognitive control task did not impair performance despite participants reporting to find the notification highly distracting. In a similar vein, other studies found that people are not good estimators of their own smartphone or internet use (Ellis et al., 2019; Scharkow, 2016). Taken together, these findings suggest that there might be a gap between what people themselves report about the distractions of their smartphones and the actual behavioral impairment these devices exert on them.

Our null findings have another potential explanation. In our experiment, we presented app icons in complete isolation from their usual context, which may have reduced their reward value altogether. Drawing from the theory of grounded cognition (Barsalou, 2009), experiences are stored in people's mind within a complex structure of sensory input, cognitions, affective states, and situational cues. It is possible that drawing a specific cue from such a rich, situated experience reduces the value of that cue. It is likely that it is only in their real-life context that smartphone apps represent social reward, because context is part of the reward value-association. Consequently, participants might explicitly evaluate app icons as rewarding if there is no time pressure and they can imagine the icons within the context of their own phones, as in Study 2. In contrast, when these app icons get isolated from their real context in the lab, they may not affect attention (M. Best & Papies, 2017). Thus, the lab situation may not be appropriate for behaviorally measuring reward associations with smartphone apps. Supporting this reasoning, it has been shown that other types of rewards, such as food, are often stored in memory in terms of situations, for instance, where people eat them (e.g., popcorn is associated with cinema) and whom people eat those foods with (e.g., family event; Papies, 2013). Future research could address this issue by measuring smartphone distractions in their natural context, for example, on people's own smartphones.

In sum, our results suggest that social app icons do not impair visual search. Given that social apps are rated as more rewarding outside of the lab, the lack of an effect of these apps on attention could be due to a loss of associated reward when taken out of context or insufficient absolute reward levels.

To our surprise, we found that participants who had locked away their phone an hour prior to the experiment were overall faster on the visual search task than participants who had not locked their phone away. However, this effect should be interpreted with caution for several reasons. First, we did not predict nor preregister a main effect of the deprivation condition. As such, the effect needs to be regarded as exploratory until independently replicated. Second, the p-value for the main effect of deprivation was extremely close to the alpha-level and many scholars argue that p-values of that size have limited evidential value (Benjamin et al., 2018). Our Bayesian analysis supports a need for caution regarding the effect, as the Bayes Factor in favor of the effect was inconclusive. Third, given the complexity to calculate power for our analysis, we cannot be certain we had sufficient power to detect a main effect. As a consequence, it is possible that our

design was underpowered for the main effect, which is problematic for several reasons. Most important, underpowered studies that yield a significant estimate will necessarily overestimate the true effect size (Vasishth, Mertzen, Jäger, & Gelman, 2018). Thus, we can only reiterate that the surprising main effect of deprivation requires high-powered, preregistered replication.

Last, we cannot exclusively attribute the effect to phone deprivation, as there was more than one difference between the non-deprived and the deprived group. Whereas participants in the non-deprived group came to the lab and immediately did the visual search task, participants in the deprived group came to the lab one hour earlier to lock away their phone and were free to do as they pleased during the one hour of deprivation, except for checking social media. Consequently, those deprived participants had one more contact point with the researchers and there was no control over what they did during the hour of deprivation. This difference might present an alternative explanation for the main effect: For instance, all participants were informed in the recruitment text for Study 1 that some of them might have to lock away their phone. Thus, at the point of locking away their phones participants in the deprived group could easily deduce that they were in the experimental condition. This knowledge might have induced reactance in the form of motivation to show that they could still perform well without their phones. Future research employing such a deprivation manipulation should consider having all participants come to the lab an hour prior to the task.

That being said, if we assume that the effect reflects a true difference between the conditions, regardless of possible confounding factors, it is plausible that locking participants' smartphones away increased their motivation, which resulted in better performance. The idea that participants could get their smartphone back and access its social rewards when they were done with the task may have motivated them to perform faster. In other words, being able to check their social media and their messages after 1.5 hours may have acted as an incentive that they could receive at the end of the experiment (Aarts, Custers, & Veltkamp, 2008). In line with such a view, participants in the deprivation condition indeed reported a higher urge to check their phones. Interestingly, this improvement in speed did not come at the cost of performance, as accuracy was almost identical across the two conditions. Such an interpretation corroborates the well-established idea that incentives boost cognitive performance (Botvinick & Braver, 2015). Another account suggests that removing the smartphone as a distractor may have resulted in better performance, as smartphones have been shown to impede attention (Stothart et al., 2015; Thornton et al., 2014; Ward et al., 2017).

Regardless of whether our data support a positive effect of deprivation or no effect, at the very least our findings stand in contrast to previous experiments, where participants who were separated from their phones performed worse on cognitive tasks (Hartanto & Yang, 2016). These studies argue that smartphone separation increases anxiety (Cheever et al., 2014), which in turn leads to worse task performance. In light of these findings, our results are quite surprising and, when replicated, may have important implications. On

the one hand, this apparent discrepancy might result from the difference in manipulations between studies. Whereas previous work deprived students of their phones strictly during the tasks, participants in our experiment were deprived both before and during the task. On the other hand, previous work also showed that anxiety due to phone separation rose with time (Cheever et al., 2014); if anything, increasing the deprivation duration should have had an even stronger negative effect on performance. Even taking into account our methodological concerns surrounding the effect of deprivation, there is a clear need for more research addressing this inconsistency. Doing so is of importance, as many policy makers, for example in schools, base their policies on findings demonstrating a detrimental effect of smartphones.

Overall, according to our findings, app icons are perceived as rewarding, but they do not capture attention and therefore do not distract participants from their task. Moreover, being deprived of access to these apps might not always be detrimental. However, these conclusions are constrained to our specific study design, which is subject to several possible alternative explanations. Our findings highlight that investigating media effects is complex and requires thorough designs that take the context of smartphone stimuli into account. As such, we believe that the current inconsistencies in the literature warrant more highly powered, preregistered research before making recommendations to policy makers.

Simple Motor Responses Change Behavior Through Changes in Explicit Liking: Influencing Preferences for Smartphone Apps

This chapter is currently under revision as Johannes, N., Buijzen, M., & Veling, H. (under revision). Simple motor responses change behavior through changes in explicit liking: Influencing preferences for smartphone apps.



ABSTRACT

Human behavior can be classified into two basic categories: execution of responses and withholding responses. This classification is used in go/no-go training, where people respond to some objects and withhold their responses to other objects. Despite its simplicity, there is now substantial evidence that such training is powerful in changing human behavior toward such objects. Yet, the underlying mechanism for this behavior change is unclear. Contrary to the popular idea that go/no-go training changes behavior by strengthening inhibitory control, we propose that go/no-go training changes behavior via changes in explicit liking of objects. To date, research has not obtained convincing evidence for this latter hypothesis. In two preregistered experiments with established experimental procedures, we show that go/no-go training influences explicit liking for smartphone apps (Experiment 1 and 2), and that this liking mediates the effect of the training on consequential choices for using these apps (Experiment 2). By demonstrating changes in explicit evaluations as a possible underlying mechanism, the results shed new light on how motor response training influences behavior. This knowledge can inform development of more effective applied motor response training procedures and raises new theoretical questions on the relation between motor responses and affect.

Human behavior toward attractive objects is difficult to change. For instance, modifying food choices is famously difficult, as is reducing choices for alcoholic beverages to promote health or well-being (Marteau, Hollands, & Fletcher, 2012). Yet, in recent years there is accumulating evidence that so-called motor response training, where people execute responses to some attractive objects and not to other objects, can have profound and long-lasting effects on behavior toward these attractive objects (Aulbach, Knittle, & Haukkala, 2019; Jones et al., 2016; Schonberg et al., 2014; Turton, Bruidegom, Cardi, Hirsch, & Treasure, 2016). We know little, however, of the mechanism underlying this behavior change. Here we propose these changes in behavior for trained objects can occur as a result of changes in explicit liking of the objects.

Despite its variety, human behavior can be divided into two basic categories: execution of responses and inhibition of responses (Guitart-Masip, Duzel, Dolan, & Dayan, 2014). This basic categorization is used in one version of motor response training, called go/no-go training (GNG). Participants execute simple motor responses to images of some objects when a go cue is presented (go objects) and withhold motor responses to other objects when a no-go cue is presented (no-go objects). When these objects are foods or beverages, participants tend to consume less of no-go objects than go objects, and choose go objects over no-go objects for consumption (Allom et al., 2016; Jones et al., 2016), although some of these findings may be subject to selective reporting (Carbine & Larson, 2019).

How can effects of GNG on behavior be explained? GNG is often presented as a training of inhibitory control (e.g., Allom et al., 2016; Forman et al., 2019; Jones et al., 2016), perhaps because the training was first introduced as such (Houben & Jansen, 2011). According to this account, repeatedly not responding to attractive objects serves as an exercise to train the brain to become better at executing control in order to resist temptations (Houben & Jansen, 2011). There are, however, problems with this account. First, to the best of our knowledge, there is no evidence that GNG can improve inhibitory control to such an extent that training effects can be observed for behavior (Enge et al., 2014); in addition, evidence for the effectiveness of other inhibitory control trainings, such as the stop-signal training, is mixed at best (Beauchamp, Kahn, & Berkman, 2016; Inzlicht & Berkman, 2015). Second, strengthening inhibitory control is assumed to be a long-term process, yet previous research often observes effects of GNG after a single training session (e.g., Z. Chen et al., 2019). Third, it is questionable whether executive control can be trained with such a simple training procedure (for an elaborate discussion see Veling et al., 2017). Despite these limitations, recent studies continue to present GNG as an inhibitory control training (e.g., Carbine & Larson, 2019; Forman et al., 2019), likely because there is a lack of convincing evidence for alternative explanations of the effects of GNG.

Instead of conceptualizing GNG as a training of inhibitory control, two accounts propose alternative mechanisms for GNG effects on behavior. The first alternative account posits that participants learn to associate go objects with responding and no-go objects with stopping (M. Best, Lawrence, Logan, McLaren, & Verbruggen, 2016; Verbruggen & Logan, 2008a). According to this stimulus-response conditioning account, participants

avoid choosing no-go objects because they have strongly associated these objects with inhibition, thus automatically triggering motor inhibition (Veling et al., 2017). Such associations may subsequently influence choices for go over no-go objects; for instance, Chen et al. (2019) discuss how stimulus-stop associations may explain the effects of GNG on food choice. The stimulus-response conditioning account is closely related to the popular idea that motor response training changes implicit S-R associations that underlie habits (e.g., Aulbach et al., 2019). However, these implicit stimulus-stop associations may depend on participants' knowledge of stimulus-stop contingencies (M. Best et al., 2016).

The second alternative account posits that behavior change through GNG occurs via changes in explicit evaluations of objects. According to this stimulus-affect conditioning account, participants learn to associate go-objects with positive affect elicited by responding, and no-go objects with negative affect elicited by not responding. Indeed, GNG has been shown to lower explicit liking of no-go objects compared to go objects or untrained objects for a variety of stimuli such as abstract art-like shapes (Clancy et al., 2019), food (Chen, Veling, Dijksterhuis, & Holland, 2016), cigarettes (Scholten et al., 2019), and erotic images (Driscoll et al., 2018). Such a transfer of affect from a response to an object can be explained by operant evaluative conditioning (De Houwer, 2007; Eder, Krishna, & Van Dessel, 2019).

During operant evaluative conditioning, valence from a conditioned response transfers to an unrelated stimulus. In previous work, when a response was followed by a positive or negative outcome ($R \rightarrow O$), the response became associated with valence through operant conditioning (e.g., R_{pos}). Next, by pairing another, unrelated set of neutral stimuli with this response ($S \rightarrow R_{\text{pos}}$), valence transferred from the response to the new set of stimuli (Blask, Frings, & Walther, 2016; Eder et al., 2019). When applied to GNG, people may have learned that responding is associated with positive outcomes and withholding responses with negative outcomes (Clancy et al., 2019; Guitart-Masip et al., 2012). In fact, some scholars argue that the relations between go and positive affect (or reward) and no-go and negative affect (or punishment) are hardwired into the brain, because such relations were functional in our ancestral past (e.g., Guitart-Masip et al., 2014). Hence, during GNG, participants transfer the negative affect of not responding to the objects to which they did not respond, and transfer the positive affect of responding to the objects to which they did respond (Hughes, De Houwer, & Perugini, 2016). According to Eder et al. (2019), this transfer may occur through association or through inferences participants make about the object and their behavior ("It doesn't feel good to stop, and I stopped for this object, so I think I don't like this object that much", cf. Van Dessel, Hughes, & De Houwer, 2018b). Consequently, stimulus-affect conditioning, operating through operant evaluative conditioning, can account for the finding that no-go objects are devalued after GNG.

In sum, there are two possible accounts to explain effects of motor response training on behavior that do not rely on the problematic conceptualization of the training as inhibitory control exercise (see Figure 1). In the present work we aimed to test whether the

stimulus-affect conditioning account can explain effects of GNG on behavior. We focus on this account because there is a long tradition in psychology demonstrating the importance of evaluations for changing behavior (Ajzen, 1991; Sheeran et al., 2016). However, so far there is a lack of studies demonstrating that changes in evaluations as a result of a simple intervention, GNG, can impact behavior.

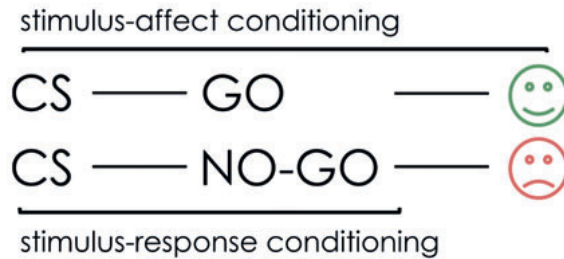


Figure 1. A graphical illustration of two possible learning mechanisms during go/no-go training. During stimulus-response conditioning, participants learn to associate go-objects (conditioned stimuli; CS) with responding, and no-go-objects (CS) with not responding. During stimulus-affect conditioning, participants learn to associate go-objects with positive affect elicited by responding, and no-go objects with negative affect elicited by not responding. Learning can be based on association or inference.

MOTOR-INDUCED STIMULUS-AFFECT CONDITIONING AS A MECHANISM TO CHANGE BEHAVIOR

It is unclear whether changing evaluations by simple motor responses is strong enough to explain behavior change in case of GNG. That is, to test changes in stimulus evaluations as a mechanism for behavior change, there is a need for evidence that changes in liking mediate the effect of the training on behavior. To date, such evidence is lacking for several reasons. First, studies demonstrating an effect of GNG on objectively measured, consequential behavior did not investigate the role of liking (Z. Chen et al., 2019). Second, the few studies measuring both a form of liking and behavior did not assess actual, consequential behavior. For instance, previous work reporting mediation relied on self-reported retrospective behavior (Houben, Havermans, Nederkoorn, & Jansen, 2012) or hypothetical choices (Veling, Aarts, & Stroebe, 2013). In addition, because these studies did not employ a priori power analysis, they might well have been underpowered to detect a mediation effect. Another study employed an indirect measurement of behavior (i.e., weight-loss), self-reported eating behavior, and explicit liking of food after GNG, but did not observe evidence for mediation (Lawrence et al., 2015). Taken together, there is no convincing evidence that GNG changes behavior through changes in object liking.

As a result, to date it remains debated whether GNG, but also other response training procedures such as approach/avoidance training, influence behavior via changes in

explicit or implicit evaluations (e.g., Aulbach et al., 2019; Veling et al., 2017). As case in point, GNG influenced implicit evaluations of alcoholic beverages, as assessed by the Implicit Association Test, but these implicit evaluations presented only a marginally significant predictor of self-reported alcohol intake (Houben et al., 2012). In a similar vein, an approach/avoidance training did influence alcohol-approach tendencies and implicit associations, but neither of these implicit measures mediated the effect of the training on treatment outcome (Wiers, Eberl, Rinck, Becker, & Lindenmeyer, 2011). Instead, recent evidence suggests that motor response training may have stronger effects on explicit rather than implicit evaluations (Van Dessel, Hughes, & De Houwer, 2018a). Hence, in addition to employing self-reported or hypothetical measures of behavior, previous studies likely did not observe evidence for mediation because measurements of implicit evaluations may not be suitable (and/or reliable, Jones et al., 2016) to capture changes in evaluation caused by motor response training. Therefore, the first goal of this study is to test whether motor response training in the form of GNG leads to changes in behavior via changes in explicit object evaluation.

A secondary goal of this study is to test whether effects of GNG can also be observed for an unstudied class of objects: smartphone apps. We chose smartphone apps for two reasons. First, they do not represent the same kind of primary reward as food, alcohol, cigarettes, or sex that have been the focus of investigation in the response training literature (Allom et al., 2016; Aulbach et al., 2019; Jones et al., 2016). If GNG changes behavior via a general learning mechanism such as stimulus-affect conditioning, it should change preferences for this unexplored category of objects. Second, whereas smartphone use does not appear detrimental to well-being (Ellis, 2019; Orben et al., 2019; Orben & Przybylski, 2019a), many users voice concerns about decreased productivity because of smartphone distractions (Johannes, Dora, et al., 2019). Therefore, GNG may serve as an intervention to modify preferences for smartphone apps, presenting a promising tool for those users who would like to reduce their smartphone use.

THE PRESENT RESEARCH

In Experiment 1, we aimed to establish a clear stimulus-affect conditioning effect for smartphone apps. We predicted that GNG would cause no-go smartphone apps to be liked less from pre-training to post-training, compared to both go smartphone apps and smartphone apps not used in GNG (i.e., untrained apps). We did not predict increased liking for go items as the version of GNG we employed here tends to influence evaluations of no-go rather than go objects (Z. Chen et al., 2016). In Experiment 2, we predicted that the evaluations of smartphone apps modified by GNG would fully mediate the effect of the training on consequential choices for using these apps.

EXPERIMENT 1

We preregistered our hypotheses, sampling plan, exclusion criteria, and confirmatory analysis plan (https://osf.io/tu3xw/register/5771ca429ad5a1020de2872e?view_only=cc109bcd77a64c3a9f67c203c9ea5b79), and provide access to all data and stimulus materials on the Open Science Framework (https://osf.io/7ck43/?view_only=a281dea1274f4a1981da310de0981719).

SAMPLE

Power calculations with mixed-effects models can be complicated (Brybaert & Stevens, 2018). Therefore, we used the *simr* package (Green & Macleod, 2016) to simulate power based on the data of Experiment 1 of Chen et al. (2016), which was almost identical to our design. Given the novelty of smartphone stimuli as objects and to properly power our experiment, we calculated power for 75% of the effect size they found. To detect such an effect with 80% power at $\alpha = .05$, we needed to recruit 63 participants. Note that 63 participants are also needed to achieve 95% power for a repeated-measures ANOVA assuming $\eta^2 = .117$, which is 75% of the smallest effect size reported in Chen et al for experiments similar to our design. In order to account for exclusion criteria, we recruited 70 participants. Participants were students from our institute and received credit or €10. We obtained IRB approval; all participants gave informed consent.

We had three inclusion criteria: First, we only recruited students between 18 and 30 years, as they are part of the population who display the strongest phone use (CBS, 2018). Second, because the go/no-go training required at least 30 app icons, participants needed to a) have 30 apps installed, and b) rate a minimum of 30 app icons during the pre-evaluation. Third, we only recruited iPhone users. All participants should be familiar with our stimuli; contrary to the Android OS, whose open source nature allows manufacturers to amend app icons, icons are standardized across iOS devices. This way, we could be certain the app icons participants rated were identical to those they know from their own phones. None of the participants fulfilled our preregistered exclusion criteria of (1) 85% accuracy or lower on the go/no-go task, where exceeding the response window counted as incorrect, and (2) a mean on the pre-evaluation of lower than -50 across all conditions. Thus, our final sample was $N = 70$ ($M_{age} = 22.20$, $SD_{age} = 2.54$, 55 female).

DESIGN

We employed a 3 (Condition: go vs. no-go vs. untrained) by 2 (Time: pre vs. post) within design with app icon evaluation as dependent variable.

PROCEDURE

DEPRIVATION PERIOD

Research on GNG or cued-approach training and food evaluations or preferences usually asks participants to fast before the experiment (Z. Chen et al., 2016; Zoltak, Veling, Chen, & Holland, 2018) to reduce between-participant differences in hunger levels which may impact food ratings. In addition, fasting will ensure the food items are on average at least somewhat appealing so that no-go devaluation can occur. We emulated this procedure for smartphone apps: Participants came to the lab one hour before the experiment and locked their phones away. This way, we minimized between-participants differences in how recently participants had used their phones right before the experiment started. By minimizing this difference, we reduced the probability that app icon ratings might be strongly influenced by one specific interaction with an app right before the start of the experiment. In addition, an hour of deprivation has shown to increase motivation to use one's phone (Johannes, Dora, et al., 2019), ensuring that apps were perceived as attractive for the rating task. After handing in their phones, participants were free to go about their day, but we instructed them not to use any of their iPhone applications on a laptop or tablet. After one hour, they returned and did the experiment. Participants received their phones back after the study.

MATERIALS

For the Go/No-Go training, we needed ratings of 30 app icons. Because we could not be sure which apps participants were familiar with, we presented them with a large selection of pre-installed and popular apps. Specifically, we selected 43 app icons that are preinstalled on each iPhone with iOS 11; in addition, we selected the top 100 list of free apps on iTunes (<https://www.apple.com/nl/itunes/charts/free-apps/>, April 3rd, 2018). Icons were presented at the center of the screen against a white background. The task was programmed with Python 2.7 (Python Core Team, 2018), using PsychoPy (Peirce, 2007).

PRE-TRAINING EVALUATIONS

Participants were instructed that the first part was about what makes an app icon look attractive to users. Thus, they rated the 143 app icons on the question "How attractive does this app icon look to you?" on a Visual Analogue Scale, ranging from -100 (*not at all*) to 100 (*extremely*). We chose this question wording because it was closest to the validated question used in previous studies on rating food pictures (e.g., Z. Chen et al., 2016). Item order was randomized. If participants did not know an app, defined as not being familiar or never having used the app, they could press the "D" button to skip the rating. All 143 apps received at least one rating and participants were familiar with a large number of apps ($M = 65.81$, $SD = 23.78$). The highest rated app (rated by at least seven participants, 10% of the sample) was *WhatsApp* ($M = 61.43$, $SD = 42.89$); the lowest rated app was *Egg* ($M = -35.44$, $SD = 39.43$).

CONDITION ASSIGNMENT

After the pre-rating, the experimental Python program rank-ordered the app icons from highest to lowest. It then repeatedly assigned the three conditions (go, no-go, untrained) from highest to lowest rated app icon for the 30 highest rated apps. To minimize pre-rating differences between conditions, item assignment was mirrored across the 30 apps starting from the top ranked icon (e.g., go, no-go, untrained, untrained, no-go, go etc.). This order was counterbalanced across participants. There were ten app icons per condition.

GO/NO-GO TRAINING

After the pre-training evaluations, participants were informed that they would do an attention task and we were interested in how well people can focus their attention while looking at different apps. During the go/no-go training, each trial began with a single app icon in the middle of the screen. After 100ms, participants heard one of two tones via headphones for 300ms. The tones were at the 1000 Hz and the 400 Hz frequencies and served as go or no-go cue. Which tone served as which cue was fully crossed with the condition assignment. For go cues, participants had to press the "B" key as quickly as possible. For no-go cues, they were instructed to not press any key. To rule out that participants were affected by reduced exposure time if the app icon disappeared after the "B" press, go and no-go icons both stayed on screen for 1000ms. Intertrial interval was random in steps of 100ms for each trial and ranged between 1000ms and 1500ms.

Participants first received a practice block of 20 trials. Icons for the practice trials were taken from the bottom of the list that rank-ordered all 143 app icons, from which the program also selected the 30 highest rated icons for the training. Thus, icons used for practice trials either received no or very low ratings in the pre-training evaluations. These icons were then randomly assigned to the go or no-go condition. During practice trials, participants received error feedback. After the practice block, participants were given the opportunity to practice again. If they chose to proceed, participants received 160 total experimental trials. The twenty icons were presented eight times; presentation order was random. After each 40 trials, participants could take a short break and received progress feedback. There was no more error feedback during the experimental block.

POST-TRAINING EVALUATIONS

After the go/no-go training, participants did the same rating task again, but this time only for the 30 selected icons. They were instructed to rate how attractive each app icon was as if it were the first time they saw it.

RESULTS

We used R (version 3.5.0; R Core Team, 2018) for all analyses. There was strong evidence that our condition assignment was successful in creating conditions that were matched on pre-training ratings; a Bayesian repeated-measures ANOVA with the *anovaBF* command

(BayesFactor package, version 0.9.12-2; Morey & Rouder, 2015) with the default priors of the function indicated a Bayes Factor of 173 in favor of a model with no differences between the three conditions. Evaluations of the icons decreased from pre-training ($M = 53.15$, $SD = 21.47$) to post-training ($M = 32.50$, $SD = 19.72$). This is in line with previous work, and generally interpreted as regression to the mean (Z. Chen et al., 2016). Similar to previous work, all participants were highly accurate during the go/no-go task ($M = 98.4\%$, $SD = 1.6\%$). The mean reaction time on correct experimental go trials was 399 ms ($SD = 56$ ms).

CONFIRMATORY ANALYSES

Following our preregistered analysis plan, we calculated difference scores between post-training evaluations and pre-training evaluations (post minus pre), where lower scores indicate stronger devaluation ($M = -20.65$, $SD = 19.00$). We tested the effect of training on the difference score with a linear mixed-effects model using the *lmer* function (lme4 package, version 1.1-17; Bates et al., 2015). Following recommendations on best practices for mixed-effects models to avoid inflated Type I error, we employed a maximal random effects structure (Barr et al., 2013). Because we had two grouping factors, we modeled two random intercepts, one for participant and one for icon in order to account for the nested nature of the data. In addition, we modeled condition (i.e., go vs. no-go vs. untrained) as a fixed effect and as a random slope varying across participants and icons.

The model converged without warnings. In line with our preregistration, to obtain p-values, we computed bootstrapped Likelihood Ratio Tests using the *mixed* function (afex package, version 0.20-2; Singmann et al., 2018), which in turn calls the function *PBmodcomp* (*pbkrtest* package, version 0.4-7 Halekoh & Højsgaard, 2014). All tests relied on 10,000 bootstraps. The effect of condition on the difference score was significant, $PBtest = 16.93$, $p < .001$. To obtain an approximation of the effect size, we squared the correlation between observed and fitted values, $R^2 = .38$. Further, we called the *r.squaredGLMM* function to obtain *Pseudo R²* for mixed-models (Barton, 2018), which yielded an estimate of .009 for the variance explained by fixed factors, and .34 for the variance explained by both fixed and random factors.

Following our preregistration, we conducted follow-up models to investigate the pairwise comparisons, see Figure 2. As predicted, no-go stimuli ($M = -25.26$, $SD = 22.09$) had a significantly lower difference score than go stimuli ($M = -16.90$, $SD = 18.50$), $PBtest = 13.95$, $p < .001$. No-go stimuli also displayed a significantly lower difference score than untrained stimuli ($M = -19.78$, $SD = 22.20$), $PBtest = 12.35$, $p < .001$. This specific pattern of results is known as the no-go devaluation effect.

EXPLORATORY ANALYSES

The main effect of condition also aligned with a Bayesian repeated-measures ANOVA with the standard Cauchy prior, with condition as within-subjects predictor and difference scores as outcome, which displayed very convincing evidence (Lee & Wagenmakers, 2013) in favor of the model with condition as predictor compared to a null model, $BF = 10,360$.

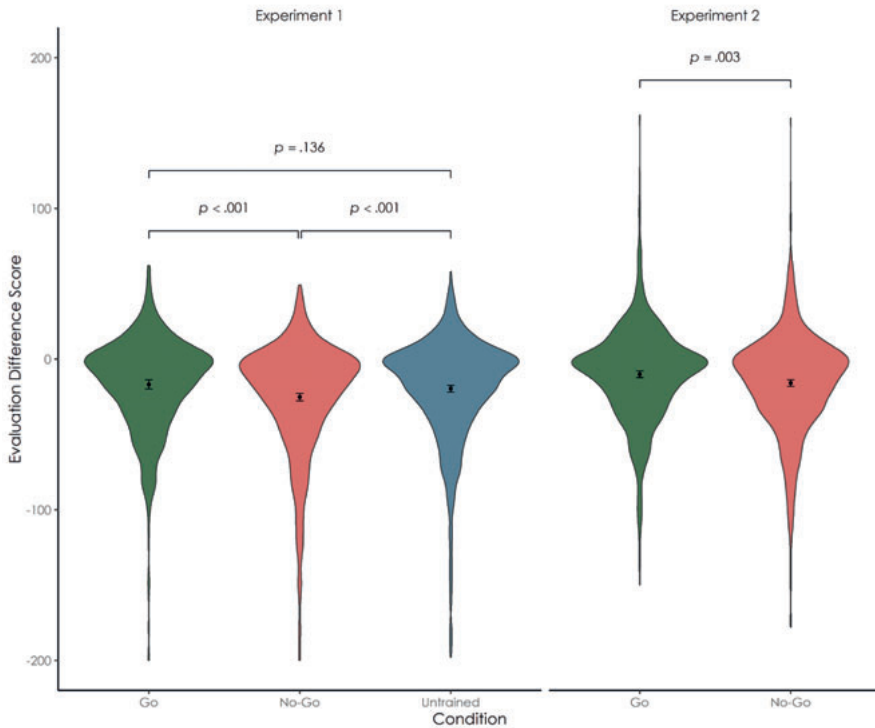


Figure 2. Violin plots of the evaluations for both experiments. Black dots in the violins represent the mean; bars of these points represent the 95% CI of the within-subjects standard error (Morey, 2008). The difference score was calculated by subtracting pre-training evaluations from post-training evaluations.

Follow-up paired t-tests with the standard Cauchy prior displayed strong evidence in favor of a difference in difference scores between no-go and untrained stimuli ($BF_{10} = 45.39$), and between no-go and go ($BF_{10} = 139$).

Our design typically leads to devaluation of no-go apps, but not to increased evaluation of go apps (Z. Chen et al., 2016). In line with previous research, although go stimuli had higher difference scores than untrained stimuli, this difference was not statistically significant, $PB_{test} = 2.12$, $p = .136$. A Bayesian paired t-test on difference scores showed that the data provided anecdotal (Lee & Wagenmakers, 2013) evidence for a lack of a difference between these conditions, $BF_{01} = 2.59$.

Furthermore, it is possible that the negative affect people experience is due to making errors during the GNG. As such, errors could explain the devaluation effect we found. To rule this possibility out, we ran the same model again excluding all items from the evaluation tasks on which participants made at least one error during the GNG. Excluding these icons did not explain the effect, as the all model parameters remained stable; that

is, no-go apps still had a lower difference score than go and untrained apps ($p < .001$ for main effect).

Finally, as an additional robustness check, we also tested whether our results were robust to excluding possible influential participants or icons. Specifically, we removed three participants and three icons because they stood out on plots visualizing Cook's distance and DFBETAs. All tests were robust to the exclusion of these cases and all parameters remained virtually unchanged (all significant differences remained at $p < .001$).

DISCUSSION

Experiment 1 showed the predicted devaluation effect of no-go apps in line with stimulus-affect conditioning : no-go apps decreased more in evaluations than both go apps or untrained apps. Consequently, Experiment 1 provides initial evidence that motor-induced stimulus-affect conditioning can be regarded as a general learning mechanism, as the effect of the training generalized to objects that are rather different from appetitive objects used in previous research.

More important, the results demonstrated that we could proceed to test the crucial mediation hypothesis. For Experiment 2, we had three aims. First, we aimed to test whether the effect found in Experiment 1 was robust and would replicate. To that end, we predicted again that no-go apps would decrease stronger from pre-training to post-training evaluations compared to go apps. Second, we aimed to demonstrate the effect of the training on preferences. Namely, we predicted that the probability of choosing go over no-go items for actual use would be significantly higher than 50%. Third, we aimed to test the mediation mechanism. During stimulus-affect conditioning, the affect a response elicits transfers to objects, which is responsible for changes in evaluations. This change in likings should then influence choices. When participants must choose between a go and a no-go app, we expected they would make their decision based on the difference in evaluations between those two apps. Therefore, we predicted that evaluations would fully mediate the effect of the training on choices.

EXPERIMENT 2

Preregistration of this experiment, specifying our hypothesis, sampling plan, exclusion criteria, and analysis plan (https://osf.io/2jm4f/?view_only=904f0b6387ae4717bbb8bddc430ac2e3), as well as all data, analysis scripts, and materials can be found on the Open Science Framework project of this manuscript (https://osf.io/7ck43/?view_only=a281dea1274f4a1981da310de0981719).

SAMPLE

According to our power simulation for Experiment 1, we needed 63 participants to reliably detect an effect of the training on evaluations with 80% power. To detect the effect of the training on choice, we followed recent power simulations by Chen et al. (2019). They

showed that 60 participants are needed for 80% power to detect an effect of at least the meta-analytic effect size of $d = 0.50$ (Allom et al., 2016). Just like in Experiment 1, we aimed to be conservative with our sample to account for the novelty of our stimuli. Therefore, we collected a roughly 25% larger sample than would be required according to the simulations; that is, we preregistered to recruit 80 participants ($M_{age} = 22.33$, $SD_{age} = 2.20$, 57 female). Inclusion criteria were the same as in Experiment 1, except that this time participants had to have 35 rather than 30 apps installed, and needed to rate at least 32 apps during the first rating task. No one fulfilled our preregistered exclusion criteria of (1) 85% accuracy or lower on the go/no-go task, where exceeding the response window counted as incorrect, (2) a mean on the pre-evaluation of lower than -50 across all conditions, or (3) choosing one side during the choice task in 90% of the time or higher. We obtained IRB approval; all participants gave informed consent.

DESIGN

We employed a 2 (Condition: go vs. no-go) by 2 (Time: pre vs. post) within design with evaluation as dependent variable, and the same condition factor but only one measurement with choice as dependent variable.

PROCEDURE

DAY 1

The procedure on the first day was identical to Experiment 1, except for the following changes. First, we updated the Top 100 free apps on iTunes (October 2018). Second, during the pre-training evaluations, participants were instructed to only rate apps they had installed on their phones and to skip rating those apps they had not. The change in instructions was evident in two instances: not all apps received a rating this time (130 compared to all 143 in Experiment 1) and participants rated less apps on average ($M = 45.45$, $SD = 8.08$). The highest rated app (rated by at least eight participants, 10% of sample) was *Netflix* ($M = 57.67$, $SD = 35.89$); the lowest rated app was *Watch* ($M = -29.33$, $SD = 44.99$). Third, we omitted the untrained condition because a) we already established that no-go items were rated as significantly lower than both untrained and go items, and b) choices may be influenced by familiarity with the icons and untrained items receive less visual exposure than the go and no-go icons. Consequently, the Python program again rank-ordered the apps icons from highest to lowest. It then repeatedly assigned the two conditions (go, no-go) from highest to lowest app icon for the 32 highest rated apps. This condition assignment was counterbalanced.

Finally, in Experiment 1, icons were presented completely randomly across all trials during the go/no-go task, which can lead to several icons of the same condition occurring in succession. For Experiment 2, the 32 experimental items were presented twice per

block, over four blocks, and the order was randomized within each block. Thus, there was a total of 256 experimental trials.

DAY 2

CHOICE TASK CONSTRUCTION

On the second day participants received a choice task. We constructed choice trials from the 32 highest rated apps on the pre-training evaluations from the first day, which were also used during the go/no-go task. We created two kinds of choice pairs. First, experimental choice pairs consisted of choices between go and no-go icons that were matched on pre-ratings. These pairs allowed us to test whether using go apps would be preferred over using no-go apps. Second, for the purpose of validating both the evaluation task as well as the choice task, we also included filler choice pairs of two apps that differed in value, but were both of the same condition (i.e., both go or both no-go). If the ratings of the app icons and choices were meaningful to participants, participants should prefer higher rated apps over lower rated apps on these filler trials. From the rank-ordered list of 32 apps, going from highest to lowest, we divided apps into ranks in order to construct experimental and filler choice pairs: (1) eight apps for experimental choice trials high in value, (2) four apps for filler choice trials high in value, (3) eight apps for experimental choice trials medium in value, (4) four apps for filler choice trials low in value, and (5) eight apps for experimental choice trials low in value (see Figure 3).

Within each of the three experimental ranks of 8 icons (high, medium, low), go icons were always paired with no-go icons; thus, there were 16 possible combinations of go and no-go icons in each rank (4 go icons x 4 no-go icons). We obtained those combinations for each of the three experimental ranks separately to keep the value difference between go and no-go icons for each experimental choice pair relatively low. Thus, there were a total of 48 unique go/no-go pairs (i.e., 16 pairs each for high, medium, and low experimental rank). These 48 unique pairs were then presented twice in the choice task, where the position of the icons (left or right) was counterbalanced, resulting in a total of 96 experimental choice trials per participant.

Between the two filler ranks, go icons of the high value rank were always paired with go icons of the low value rank, and no-go icons of the high value rank were always paired with no-go icons of the low value rank. There were two go icons and two no-go icons in each filler rank, thus allowing for eight unique possible combinations of choice pairs within the same condition, but with different values: (2 high value go icons x 2 low value go icons) + (2 high value no-go icons x 2 low value no-go icons). Just like with experimental choice trials, we counterbalanced the position on in the choice task, thus leading up to 16 choice trials. Last, to increase power for the effect of value on choice for filler trials and to be consistent with the number of choice trials employed in previous research (Schonberg et al., 2014), we doubled the number of filler trials once more. That is, there were 32 filler choice trials in total per participant.

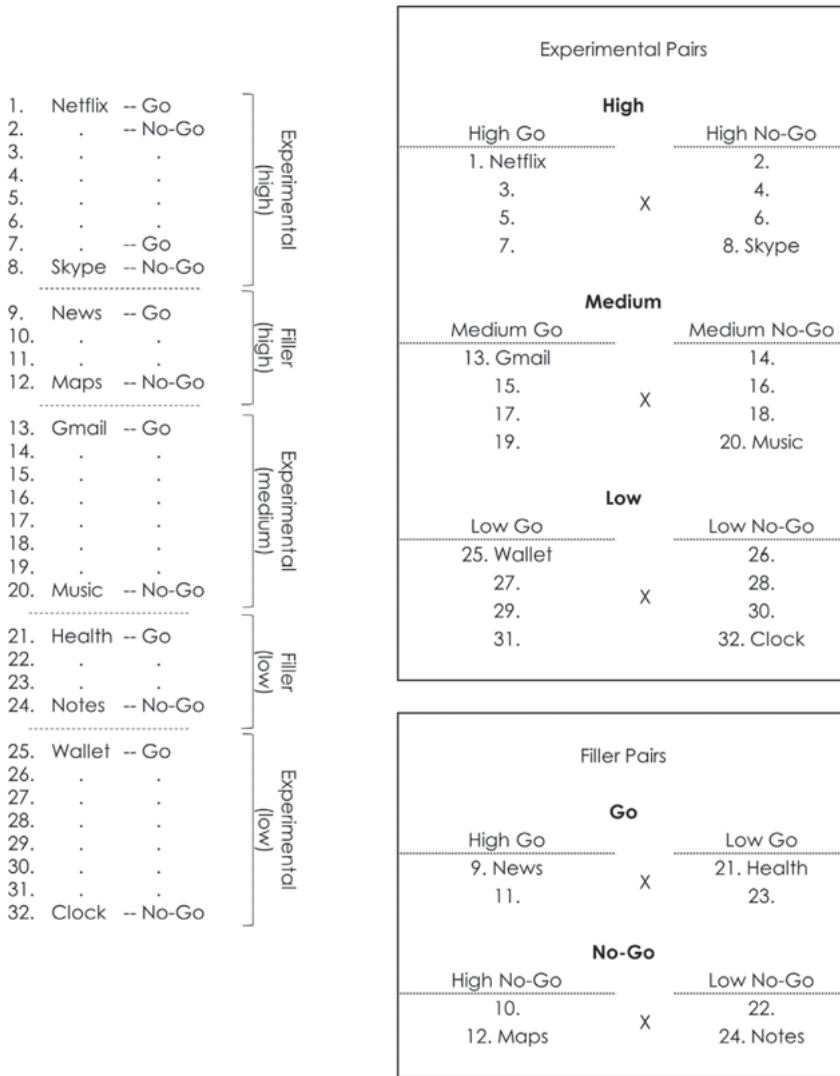


Figure 3. An illustration of how choice trials were constructed. On the left is an exemplary list of the 32 highest rated apps, ordered from highest pre-training evaluation to lowest. The assignment of the condition (go vs. no-go) was counterbalanced across participants. The 32 highest apps were then divided into different ranks, three experimental and two filler ranks. The two boxes on the right visualize how choice trials were constructed from these ranks.

Experimental blocks served to test the effect of condition, whereas filler blocks served as a check whether participants would indeed choose high over low value in the majority of cases. The total 128 choice trials (96 experimental + 32 filler) were presented in two blocks without a break. Each block contained half of the experimental and half of the filler trials, counterbalancing the position of each app icon. Before the experimental block, participants practiced the choice task with 16 choices between icons that were not used in the training.

CHOICE TASK PROCEDURE

Before doing the choice task, participants again locked their phones away for one hour. Apart from the reasons described earlier, this time the deprivation served an additional purpose. That is, participants made consequential choices during the choice task about which app they could use for a short while during the experimental session. Previous research showed that an hour of deprivation makes participants motivated to use their phones (Johannes, Dora, et al., 2019), which made our choice task more relevant to them. When the choice task is employed in food research, participants learn that the program will pick a random trial in the end; whatever participants chose on this trial is the food they receive (Krajbich, Armel, & Rangel, 2010). This way, choices are consequential for participants. Hence, analogue to food research, we instructed participants that the program would randomly pick a trial at the end of the choice task; whatever app participants chose on that trial was the app they were allowed to use for three minutes before we locked their phones away again for half an hour. This second deprivation phase was intended to make the choices meaningful for participants. Without that second deprivation phase it would not have mattered which trial the program picked, as participants could have just used whichever app they felt like after the experiment ended. Thus, participants locked their phones away for an additional half an hour after using the app for three minutes that they chose on the trial randomly selected by the program.

After the first hour of deprivation, participants came back to the lab and did the choice task. On each trial, participants chose between two apps that were presented side by side by pressing the "U" or "I" keys. Participants had to make that choice within 1500ms. If they chose an app within the response window, their choice was confirmed by a yellow frame surrounding the app for 500ms. If they did not make the choice in time, they were presented with feedback ("Choose faster!") and the choice trial was presented again at the end of the block. Participants rarely exceeded the response window (1.22%). Intertrial interval varied randomly between 1000ms and 2000ms in steps of 100ms.

RESULTS

EFFECTS ON EVALUATION

There was no evidence of a difference in ratings between go and no-go items at the pre-training evaluations ($BF_{01} = 23.02$). Once more we observed regression to the mean from pre-training ($M = 36.72$, $SD = 22.64$) to post-training evaluations ($M = 23.66$, $SD = 21.24$). Accuracy was high ($M = 98.5\%$, $SD = 1.3\%$) and mean reaction time on correct experimental go trials was 471ms ($SD = 54$ ms).

CONFIRMATORY

Similar to Experiment 1, we preregistered to test the effect of the training on the difference score between pre-training and post-training ($M = -13.07$, $SD = 12.28$), see Figure 2. Again, we employed a linear mixed-effects model with a maximal random effects structure. The initial model failed to converge, most likely because there were 18 apps that received only one rating. Estimating a random slope for the difference between the go and no-go condition for each app requires the app to have a rating for each condition. With only one evaluation, the model cannot estimate a random slope. Consequently, it was necessary to group those apps into an “other” category; the model then treated the 18 evaluations as coming from the same group (i.e., app). Afterwards, the model converged without warnings. In line with our preregistration, we followed recommendations by Luke (2017) and obtained p-values with Satterthwaite approximation for degrees of freedom. The difference between go items ($M = -10.17$, $SD = 14.07$) and no-go items ($M = -15.97$, $SD = 15.34$) was significant, $F(1, 66.49) = 9.67$, $p = .003$. Squaring the correlation between observed and fitted values yielded $R^2 = .22$. *Pseudo* R^2 for the variance explained by the fixed factors was .008, and .17 for the variance explained by both fixed and random factors.

EXPLORATORY

The main effect of condition aligned with a Bayesian repeated-measures ANOVA, $BF_{10} = 4,472$. Again, excluding icons on which participants committed an error did not influence the model ($p = .004$). In addition, we excluded two participants who had a large influence on the estimates, as indicated by Cook’s distance and DFBETAs, and ran the same mixed-effects model again. The effect was robust to excluding outliers ($p = .008$). Last, we ran the same model again without those 18 apps that we grouped into the “other” category to ensure the effect was not driven by those cases. The effect was robust to the exclusion of this category ($p = .003$).

EFFECTS ON CHOICE

Participants chose go items over no-go items on 54.8% of experimental trials ($SD = 12.3\%$). This percentage was similar across the three value ranks: 53.5% for choices high in value, 56.0% for choices medium in value, and 55.0% for choices low in value.

CONFIRMATORY

To test whether the overall percentage of choosing go items over no-go items was different from 50%, we ran a generalized mixed-effects intercept-only model with a random intercept per participant. Please note that we deviated from the preregistration here: we preregistered to obtain the p-value with Satterthwaite approximation for degrees of freedom which is not possible for a generalized intercept-only model. Instead, we report the p-value based on Wald's test statistic. Because Wald's test statistic can be problematic, we also report the 95% confidence interval obtained with the profile method. The fixed intercept was significantly different from 0, estimate = .21, $SE = .059$, $z = 3.50$, $p < .001$, indicating that participants chose go items above chance level, $OR = 1.23$, $95\%CI = [1.09, 1.38]$. That is, participants had 1.23 times higher odds of choosing go items than choosing no-go items.

EXPLORATORY

Participants made valid choices in the choice task: On filler trials, when both app icons on a choice trial were go items or both app icons were no-go items, but they differed in value, participants chose the higher valued icon 66.9% of the time ($SD = 18.8\%$, Figure 4). A generalized mixed-effects intercept-only model with a random intercept per participant shows that participants indeed chose high-value apps above chance level, estimate = 0.82, $SE = .11$, $z = 7.46$, $p < .001$, $OR = 2.27$, $95\%CI = [1.83, 2.83]$.

As a robustness check, we also analyzed the choices for go versus no-go items with a Bayesian one-sample t-test comparing the mean proportion of choices for go items against 50%. This test indicated strong evidence that choices were different from chance level, $BF_{01} = 30.86$.

There were no visual or formal outliers to test for robustness of the model. We were also interested whether the effect of the training would differ for different choice reaction times, as previous research has observed that the choice effect becomes weaker the more time participants take (Z. Chen et al., 2019). However, although the coefficient was negative, reaction times were not a significant predictor of app choice $\chi^2 = .2.23$, $p = .135$.

MEDIATION

Last, we tested whether evaluations mediated the effect of the training on choice. However, we could not conduct a direct test of the mediation, as the evaluation data and the choice data had different structures. For the evaluations, each case (i.e., app icon) had one condition assigned. For the choices, each case (i.e., choice trial) presented two apps and thus two conditions alongside each other. As a consequence, we could not assess mediation in a single statistical test. Instead, we preregistered to employ the approach of joint significance: If the effect of the independent variable (i.e., the training) on the mediator (i.e., the evaluations) is significant, and the effect of the mediator on the outcome (i.e., choosing go over no-go) is significant, it follows that the indirect effect is likely non-zero (Kenny, Kashy, & Bolger, 1998). Although joint significance does not provide a single

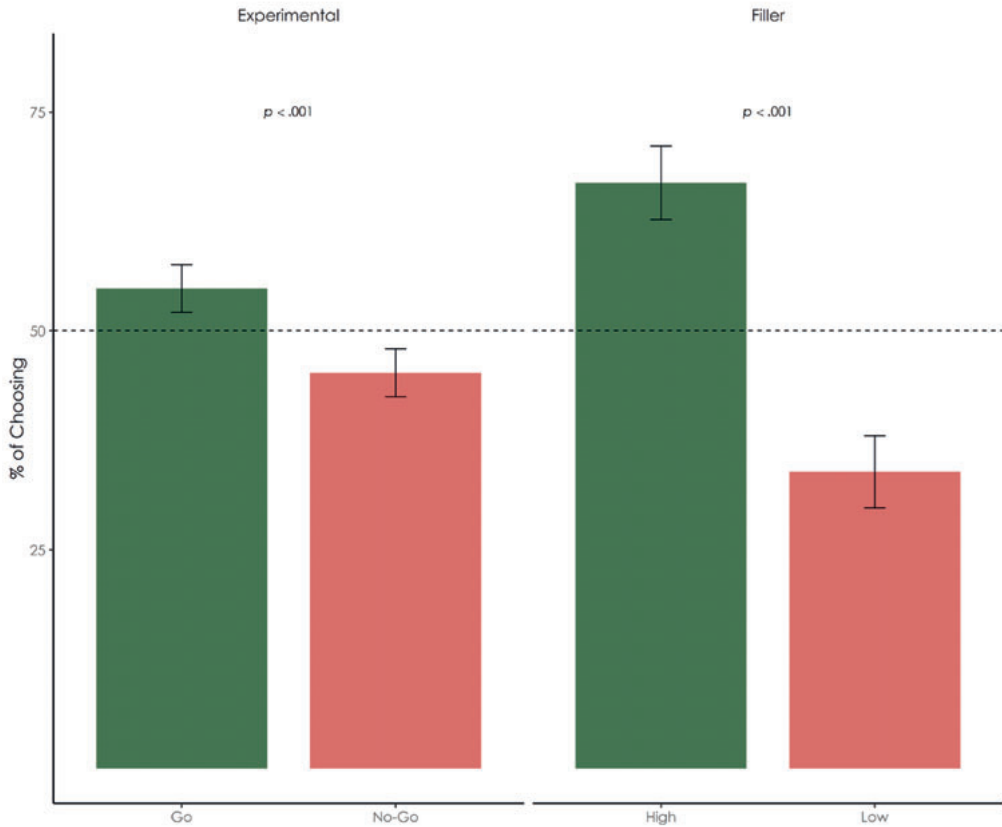


Figure 4. The left side of the graph shows the mean probability of choosing go and no-go apps in experimental trials. The right side of the graph shows the mean probability of choosing high-value and low-value apps in filler trials. P-values reflect the overall tests of choosing go over no-go (left) and high-value over low-value (right) against chance level (50%, dashed line). Error bars reflect the 95% CI of the standard error.

estimate of the indirect effect size, it performs well in assessing mediation compared to other procedures such as bootstrapping (Hayes & Scharkow, 2013). Consequently, we first tested whether the training would influence evaluations; afterwards, we assessed whether the evaluations predicted choice.

CONFIRMATORY

To test the first path of the mediation model, we estimated a maximal mixed-effects model with the condition of app icon as predictor and the post-training evaluations as outcome. We preregistered to use the post-training evaluations as mediator because a) we already established an effect of the training on the evaluations controlling for the pre-training evaluations, and b) we wanted to avoid using a difference score (go over no-go)

of a difference score (post minus pre) to aid interpretation. As predicted, no-go items ($M = 20.75$, $SD = 23.21$) were rated significantly lower than go items ($M = 26.56$, $SD = 22.45$), $F(1, 78.91) = 10.38$, $p = .002$.

To test the second path of the mediation model, we estimated a maximal mixed-effects model with the difference score of post-training evaluations as predictor and choice of go icons as outcome. Specifically, because each trial presented two app icons, we calculated the difference between post-training evaluations of the go item and the no-go item (go minus no-go). Thus, positive values reflect how much higher participants evaluated the go item compared to the no-go item on that specific choice trial. In line with the effect of the training on post-training evaluations, go items received higher ratings than no-go items ($M = 5.77$, $SD = 19.21$). This analysis enabled us to carry out the test of joint significance: For the first path of the mediation model, we tested the effect of training condition on post-training evaluations; if those post-training evaluations predicted whether people chose go items over no-go items, we could have confidence that evaluations indeed mediated the effect of training on choice. Importantly, if the intercept in the model did not remain significantly different from 0 (i.e., chance level), whilst including evaluations as predictor, this would indicate full mediation, because the evaluations can account for all variance in choices.

Note that for the reasons described above we again had to deviate from the preregistration and obtain p-values with likelihood-ratio tests. As predicted, higher evaluations for go items compared to no-go items was a significant predictor of choosing go items over no-go items, estimate = .015, $SE = .002$, $\chi^2(1) = 54.03$, $p < .001$, OR = 1.015, 95%CI = [1.011, 1.018]. Thus, with each point that participants rated go items higher than no-go items, they had 1.015 higher odds of choosing go items over no-go items. Although these odds may seem small, evaluations were made on a VAS scale ranging from -100 to 100. To get a better understanding of the effect size, the odds ratio for the average difference between go and no-go items was $1.015^{5.77} = 1.09$. If participants showed a difference of one standard deviation in their rating of go and no-go items, they would have $1.015^{19.21} = 1.33$ higher odds of choosing go over no-go items, larger than the overall effect of the training on choice, see Figure 5.

Interestingly, the intercept was also significant, estimate = .16, $SE = .06$, $z = 2.82$, $\chi^2(1) = 7.59$, $p = .006$, OR = 1.17, 95%CI = [1.05, 1.31]. Even after accounting for the effect of evaluations, participants still had 1.17 higher odds of choosing go items over no-go items. Hence, evaluations appear to only partially mediate the effect of the training on choices.

EXPLORATORY

A Bayesian repeated-measures ANOVA indicated strong support for the main effect of condition on post-training evaluations, $BF_{10} = 411$. All analyses reported to test the mediation were robust to removal of outliers.

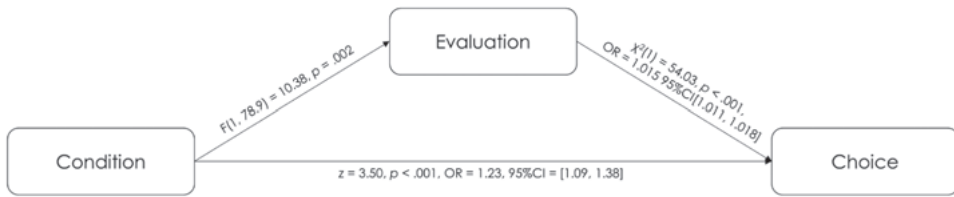


Figure 5. A visualization of the mediation model that we ran. The direct path from condition to choice represents the test of the training on choices. The path from condition to evaluation represents a test of the training on post-training evaluations. The path from evaluation to choice represents the effect of the difference in post-training evaluations between a go and a no-go app on choice.

GENERAL DISCUSSION

Motor response training tasks, such as GNG, have become popular tools to tackle behavior change (Allom et al., 2016; Jones et al., 2016; Stice, Lawrence, Kemps, & Veling, 2016; Turton et al., 2016). However, to date the mechanism underlying their effect remains debated (e.g., Aulbach et al., 2019). One reason for this debate might be the conceptualization of motor response training, as much of past research has proposed GNG affects behavior by improving inhibitory control (Houben & Jansen, 2011). However, such an understanding appears problematic and inadequate to account for the effects of GNG (Enge et al., 2014; Inzlicht & Berkman, 2015; Veling et al., 2017).

Instead, we focused on an alternative mechanism. Here, we examined stimulus-affect conditioning, induced through GNG, as a mechanism for behavior change, in line with the mechanism proposed by operant evaluative conditioning (De Houwer, 2007; Eder et al., 2019; Hughes et al., 2016). That is, we tested whether the affect that is associated with the actions of going or not going transfers to the objects on which these actions are carried out. More important, we also investigated whether such a stimulus-affect conditioning effect influences consequential behavior. In two high-powered, preregistered experiments with established measures of explicit evaluations and consequential behavior, we demonstrate robust evidence that GNG influences evaluations of smartphone apps (Experiments 1 and 2), and that this motor-induced stimulus-affect conditioning effect mediates the effect of GNG on behavior (Experiment 2).

The central role of evaluations for the effect of GNG on behavior casts doubt on accounts that understand GNG as a self-control training. In other words, improvements in inhibitory control cannot explain the changes in evaluations and their influence on behavior we observed. In fact, we believe that branding GNG and other motor response trainings as inhibitory control trainings might be inaccurate. Inhibitory control training is still the prevalent term used in the literature, yet the effect of motor response training on self-control has received little empirical or theoretical support (Enge et al., 2014; Inzlicht &

Berkman, 2015; Veling et al., 2017). It might be time to abandon using this term altogether and instead summarize these trainings under a broader umbrella term: motor response training.

Interestingly, the mediation we observed was only partial, and not full as we predicted. Specifically, even after accounting for the influence of evaluations, participants chose go over no-go apps above chance level, as indicated by the significant intercept. This effect suggests there is room for other mechanisms to further explain effects of GNG on behavior. We predicted that the affect associated with objects, expressed in evaluations, should fully drive the effect of the training on choices. Alternatively, stimulus-response conditioning does not account for changes in evaluations, but predicts that participants have learned to associate stopping with certain objects, thus preferring go over no-go objects. Choosing one object over the other would thus merely reflect a trained motor response (M. Best et al., 2016; Verbruggen & Logan, 2008a). Given that the mediation was only partial, it stands to debate whether both accounts contribute to the effect of GNG on behavior. When presented with the choice between a go object and a no-go object, participants might be guided in their decision by how they evaluate the two options as well as a conditioned motor response. Disentangling the effect of motor-induced stimulus-affect conditioning from other mechanisms presents promising avenues for research aiming to understand the mechanism behind motor response trainings.

Open questions notwithstanding, we show that the affect associated with a simple response can transfer to unrelated objects; and this association can drive behavior. Important, we assessed actual, consequential behavior, thereby extending previous research that employed self-reported or hypothetical measurements (Houben et al., 2012; Veling et al., 2013). Evaluations and behavior were entirely different measures, which rules out common method variance as an alternative explanation for mediation. Thus, motor-induced stimulus-affect conditioning appears to be a powerful learning mechanism that changes real behavior. It is noteworthy that we found effects on a previously unexamined category of objects, smartphone apps, suggesting effects of GNG may apply to a wide range of stimuli, similar to Pavlovian, evaluative, and operant conditioning (De Houwer, 2007). Not only did the effect generalize beyond appetitive stimuli used in previous research to a new class of objects, namely smartphone apps; we also observed effects of evaluations on choices one day later, indicative of the strength of the effect. We introduced a temporal order to strengthen claims of causality: evaluations succeeded the training, and choices succeeded evaluations. Therefore, modifying liking via transfer of affect could potentially explain the findings of previous work which shows that the effect of GNG on behavior can be observed up to two weeks later (Z. Chen et al., 2019).

However, although there is strong evidence that some motor response training procedures can influence preferences for months (i.e., cue-approach training, Salomon et al., 2018; Schonberg et al., 2014), there remain open questions. First, it is unclear whether motor-induced stimulus-affect conditioning can explain findings for these other motor response training procedures as well. Second, it remains to be tested whether the changes

in evaluations and choice behavior we observed can explain long-term behavior change. Furthermore, our behavioral measure was strictly confined to choices that participants do not encounter in such a form in their everyday lives. Hence, there is a need for research testing whether decreased evaluations also have an impact on smartphone use outside the lab. This question is crucial, as current evidence is inconsistent with regard to how applicable motor response training is outside the lab. On the one hand, the effect of the training on choices in the lab appears to be restricted to intuitive, quick choices (Z. Chen et al., 2019). On the other hand, motor response training has shown to have an effect on real-life behaviors outside the lab (Jones et al., 2016; Lawrence et al., 2015).

Our findings highlight the importance of employing explicit measures of evaluation. So far, much of previous scholarship has relied on implicit measures of liking (Aulbach et al., 2019), but has not found evidence that implicit measures mediate the relation between GNG (Houben et al., 2012) or approach/avoidance training (Wiers et al., 2011) and behavior. Our findings are in line with other recent work that suggests that motor response training has stronger effects on explicit, rather than implicit evaluations, particularly when adding clear and explicit evaluative consequences of the responses (Van Dessel, Hughes, et al., 2018a). These consequences might have emphasized the natural affect associated with the responses, ultimately amplifying the effect of the training on evaluations.

The exact mechanism of the transfer of affect during GNG remains open to debate. From the perspective of operant evaluative conditioning (De Houwer, 2007; Eder et al., 2019), when a conditioned response (e.g., using a particular key is followed by a pleasant picture) is used for another set of neutral stimuli, the affect associated with that response (e.g., the key) transfers to the new set. In our case, we did not conduct the first step, as we assumed that going was naturally associated with positive affect, and not going was naturally associated with negative affect. To test whether operant evaluative conditioning is the mechanism underlying the current effects, we call for research that combines such as an operant evaluative conditioning paradigm ($R \rightarrow O, S \rightarrow R$) with a similar choice task as we employed, having participants choose between stimuli that have been positively vs. negatively conditioned during the $S \rightarrow R$ phase.

However, even such a test would not settle the question of how the transfer of affect occurs. Eder et al. (2019) discuss two possibilities. An inferential account posits that people infer their liking of objects from their own behavior (Van Dessel, Eder, & Hughes, 2018; Van Dessel, Hughes, et al., 2018b). According to this view, participants observed that they stopped for an app and stopping is unpleasant, thus negatively adjusting their evaluation of the app. Alternatively, the app, not going, and the unpleasantness of not going might all be associated in memory (Hommel, 2004), such that retrieving one part of this association (i.e., the app) automatically retrieves the other parts (i.e., the unpleasantness), thus decreasing liking for the app (for a more elaborate discussion see Eder et al., 2019). As a consequence, we call for more research to investigate a) whether an operant evaluative conditioning paradigm can influence behavior, b) whether stimulus-affect conditioning can account for effects on implicit evaluations, and c) what accounts can explain the

affect transfer resulting from stimulus-affect conditioning and whether these accounts are complementary or not reconcilable. Apart from theoretical insight, understanding whether motor response trainings have a common mechanism can inform decisions on whether to combine different response trainings in one intervention.

CONCLUSION

Motor response trainings have proven effective in changing human behavior. However, their underlying mechanism has remained unclear. According to our findings, the affect induced by simple motor responses can transfer to objects, which influences which objects people choose to use. Hence, motor-induced stimulus-affect conditioning presents a potentially powerful mechanism for behavior change.

Discussion



It has been over a decade since the first iPhone hit the market. In this time, human communication and behavior have changed dramatically. With people always online and constantly connected to others, concerns have been expressed over users glued to their smartphones, erratic and unable to focus (Carr, 2011; Turkle, 2012). Some people appear to be bothered by constant notifications to a degree that even the sheer presence of their smartphone can be distracting (Stothart et al., 2015; Thornton et al., 2014; Ward et al., 2017). *Smartphone cues*, such as the presence of a smartphone or notifications, can contribute to a state of alertness, labeled *online vigilance* (Klimmt et al., 2018; Vorderer & Kohring, 2013). Importantly, many users experience smartphone cues and the resulting online vigilance as distracting and bothersome. Moreover, these concerns highlight the need for research testing ways to reduce the appeal of smartphone cues. In this dissertation, we had the overall aim to investigate the effects of smartphone cues and online vigilance on well-being and performance.

Specifically, we had four goals. The first goal was to investigate the relation of online vigilance with well-being. We investigated whether and how online vigilance relates to well-being on the trait (Chapter 2) and on the state level (Chapter 3). The second goal was to test whether smartphone cues interfere with the executive control functions of inhibition (Chapter 4) and working memory (Chapter 5). The third goal was to investigate whether online vigilance can be explained by the rewarding nature of smartphones (Chapter 6). The fourth goal was to test whether we can reduce people's preferences for smartphone cues (Chapter 7). In this final chapter, we discuss outcomes of these goals.

GOAL 1: DOES ONLINE VIGILANCE RELATE TO WELL-BEING?

The first central question was whether and how online vigilance is related to well-being. On the one hand, users may experience the constant connection to others as assuring, signaling social support (Antheunis et al., 2015; Domahidi, 2018; Trepte et al., 2015). On the other hand, the mobility of smartphones guarantees a constant stream of smartphone cues (Vorderer & Kohring, 2013), contributing to a feeling of alertness and mental preoccupation (Klimmt et al., 2018; Reinecke et al., 2018). This online vigilance may make users feel overwhelmed; for instance, such mental preoccupation with constant communication demands of the online sphere can manifest in feelings of entrapment and stress (Halfmann & Rieger, 2019; Hall & Baym, 2012; Reinecke, Aufenanger, et al., 2017). Perhaps not surprising, many users experience constant connectedness as bothersome (Mihailidis, 2014; Näsi & Koivusilta, 2013).

In Chapter 2, we investigated whether this perception is reflected in empirical evidence (Johannes et al., 2018). We reasoned that online vigilance would be detrimental to well-being if it takes the form of absentmindedness. Thus, we tested whether online vigilance was indirectly related to well-being through two established indicators of

absentmindedness: mind-wandering and mindfulness. Importantly, we did not assess well-being as a uniform construct. Instead, to obtain a more complete picture we followed current recommendations to assess both to the evaluative component (i.e., satisfaction with life) and the affective component of well-being (i.e., affective well-being; Diener et al., 2018). As predicted, online vigilance and the two well-being components were related indirectly: Online vigilance was related to the well-being outcomes through decreased mindfulness, but not directly. That means we observed a rare case where a mediator (i.e., mindfulness) masks the relation between two constructs (i.e., online vigilance and well-being). The negative relation between online vigilance and mindfulness and the positive relations between mindfulness and the well-being outcomes explain the small direct correlations between online vigilance and the well-being outcomes.

This mediation suggests a crucial role for mindfulness. Theoretically, a constant preoccupation with the online sphere might not be directly linked to lower well-being. Yet, if such preoccupation is an indicator of absentmindedness in the form of decreased awareness and acceptance of current thoughts, people might experience lower well-being. Supporting such a view, online vigilance was also directly and positively related to mind-wandering, another indicator of absentmindedness. If online vigilance only becomes problematic for people when it takes the form of distraction or absentmindedness, then the cognitive component of online vigilance appears to play the greatest role in explaining why users complain about being constantly alert. In other words, the salience dimension might be the most bothersome, an argument in line with experimental work that suggests that smartphone cues trigger task-irrelevant thoughts and distract from the current moment (Stothart et al., 2015; Ward et al., 2017). We tested this proposition in an exploratory analysis where salience showed to be the strongest driver of the effects we found. In fact, salience was the only dimension of online vigilance that was related to mindfulness.

This finding has several explanations. On the one hand, even if we must regard this insight with caution because it was exploratory, the central role of salience suggests that only certain aspects of online vigilance pose a threat for well-being. It is also unclear whether this threat is practically meaningful given that the effect was small and indirect. On the other hand, constant checks (i.e., monitoring) and high sensitivity to smartphone cues (i.e., reactivity) should also result in lower mindfulness, because these dimensions manifest in behavioral distractions in the form of multitasking or task-switching (Alzahabi & Becker, 2013; Jain et al., 2007; van der Schuur et al., 2015). The lack of a relation between these two dimensions and mindfulness casts doubt on the measurement of these two online vigilance dimensions. In fact, both reactivity and monitoring asked participants to assess their smartphone behavior. However, participants might not be able to accurately make such assessments (Boase & Ling, 2013; Ellis et al., 2019; Scharrow, 2016).

Furthermore, it is not clear whether self-reported traits correlate with day to day fluctuations in well-being and actual behavior (Saunders et al., 2018). Our second study, Chapter 3, addressed this issue by relying on objective behavioral measures to assess monitoring and reactivity. Objective measures present a better indicator of actual

behavior which does not suffer from biases in self-reports (e.g., Ellis et al., 2019; Vanden Abeele et al., 2013). Furthermore, all measurements were on the state level, repeated throughout the day for five weekdays. Such in-the-moment assessments provide a picture of relations in real life and enable us to disentangle within-person and between-person effects (Csikszentmihalyi & Larson, 2014; Miller, 2012). Together, these steps increased ecological validity, giving us a better understanding of smartphone cues and online vigilance.

Chapter 3 shows that the behavioral components of online vigilance, monitoring and reactivity, were not related to well-being. Contrary to our assumption in Chapter 2 that measurement error might explain why the monitoring and reactivity dimensions were not related to well-being, measuring these dimensions objectively did not result in meaningful effects. Using social apps in the past half an hour (i.e., monitoring) showed a slight negative effect on well-being at that moment, but this effect was not robust. It also did not matter for well-being how quickly participants responded to surveys (i.e., reactivity).

Interestingly, and extending findings from Chapter 2, (self-reported) salience again was the only robust predictor of well-being. The more people thought about online interactions in the past half an hour, the worse they felt in the current moment. The effect of salience aligns with previous research demonstrating that mind-wandering generally can be detrimental to well-being (Franklin et al., 2013; Killingsworth & Gilbert, 2010). Contrary to research on mind-wandering, we did not explicitly assess whether salience interrupted other tasks. This difference distinguishes salience from mind-wandering. Hence, salience can also be understood as an indicator of not being present in the current moment, in line with research showing lower mindfulness to relate to well-being negatively (Brown & Ryan, 2003; Brown, Ryan, & Creswell, 2007).

In Chapter 2, with measurement on the trait level, the quality of people's thoughts could not be easily assessed (e.g., it seems unlikely that people can accurately recall the valence of their many thoughts). Experience sampling in Chapter 3 allowed us to capture valence of thoughts. When thoughts about online interactions were positive, that positive valence more than compensated for the negative effect of the intensity of such thoughts. Practically, this means that thinking of a pleasant online interaction might distract from the current moment; but this pleasantness outweighs the distraction. Again, this finding supports previous research showing that pleasant thoughts can increase affect (Franklin et al., 2013; Poerio et al., 2016). Equally important, the effect of valence of thoughts about online interactions was comparable to the effect of valence of thoughts about face-to-face interactions. In contrast to salience intensity, thoughts about face-to-face interactions were not negatively related to well-being. This comparison shows that permanent mental preoccupation with the online world might indeed be detrimental to well-being. However, the effect becomes negligible as long as this mental preoccupation takes a positive form, in line with accounts highlighting the potential of smartphones for social support (Domahidi, 2018; Trepte et al., 2015).

INTERIM CONCLUSION 1

Based on our findings, concerns about negative effects of online vigilance might not be justified. On the level of individual differences, if online vigilance takes the form of absentmindedness it indeed relates negatively to well-being. However, that effect is small. On the situational level, mental preoccupation with online interactions again shows a small negative relation with well-being. However, the quality of such preoccupation compensates for this negative relation, such that positive thoughts are far more important than how frequently they occur.

GOAL 2: DO SMARTPHONE CUES IMPACT PERFORMANCE?

The second central question was whether smartphone cues impact performance. In addition to worries about well-being, smartphone users express concerns about constant smartphone cues distracting them from a main task (Mihailidis, 2014; Näsi & Koivusilta, 2013; Stothart et al., 2015). Research shows that media represent the everyday temptation users have most trouble resisting (Hofmann, Vohs, et al., 2012). Because smartphone cues remind users of the rewarding features of their smartphone, they experience conflict (Bayer et al., 2015). According to such a view, notifications or even the mere presence of a smartphone impair basic cognitive functions (Chein et al., 2017; Stothart et al., 2015; Thornton et al., 2014; Ward et al., 2017).

In Chapters 4 and 5, we tested whether smartphone cues (notifications and the presence of a smartphone) would impair inhibition and working memory. The experiments were the first preregistered tests of smartphone cues and employed a sequential Bayesian sampling design, optimizing power. The studies showed a surprising discrepancy between experience and performance. In Chapter 4, participants reported much higher online vigilance and distraction when their phone received a notification or even when their phone was only on the table compared to a group without their phones. However, their behavior was not affected by these cues. All groups performed similarly on the inhibition task. Mirroring those results, in Chapter 5, participants perceived notifications as highly distracting compared to a control group, but they performed equally well on a working memory task.

The results of Chapter 4 stand in stark contrast to previous research demonstrating detrimental effects of smartphone cues and attention (Stothart et al., 2015; Thornton et al., 2014). The experiment was preregistered and well-powered, enabling us to perform a thorough test of the effect of smartphone cues. Not only were we unable to conceptually¹ replicate the effect of notifications (Stothart et al., 2015); receiving notifications meant that phones were present, likening them to the mere presence condition. Consequently, we

¹ Direct replications attempt to recreate an experimental design. Conceptual replications attempt to test the same research question with slightly amended experimental designs (Kunert, 2016).

should at least be able to replicate the effect of mere presence in both of our conditions (Thornton et al., 2014). Previous research showed notification and mere presence effects on attentional measures, possibly explaining the failure to replicate the effect. However, our inhibition measure also required attention (e.g., to detect the stop cues; Verbruggen, McLaren, & Chambers, 2014), which leads us to believe that we should have at least found some effect of notification and mere presence. Consequently, the results from Chapter 4 raise questions about the reliability of previous research findings. For instance, it is possible that people's reactions toward smartphone cues have changed in recent years. Alternatively, previous work may have been underpowered. Although low power generally means a low probability to detect a true effect, it also means that occasional significant study effects are an overestimation of the true effect size (Button et al., 2013; Vasissth et al., 2018). Therefore, we propose more (direct) replications to establish whether the effect of smartphone cues on attention is robust.

For Chapter 5, we used a working memory task, making it comparable to the measure used by Ward et al. (2017). However, our sample size was small in light of the small effects reported by Ward et al. who used a large sample to be able to detect these small effects. Because their experiment was conducted after ours, we could not rely on their effect size to conduct power analysis. Thus, our experiment was likely underpowered, signified by Bayes Factors that indicated only anecdotal evidence for the lack of an effect. If effects of smartphone cues truly exist, they might be small and limited to working memory.

Overall, we find a strong discrepancy between self-reports and actual behavior: Self-reported online vigilance and smartphone distraction were unrelated to impairments in performance. Whereas the effect on self-reported distraction is strong, the effect on behavior appears to be small at best. This discrepancy is in line with our findings regarding online vigilance and well-being. Here, only self-reported cognitions were related to well-being; behavioral indicators were not problematic. Consequently, the perception of users that smartphone cues are detrimental and that they lead to more distracting thoughts might drive evaluations of and affective well-being, but not actual behavior. In other words, one possible theoretical explanation for our findings are self-evaluation errors, in line with a previous study that found instant messaging use to affect self-reports, but not performance (Levine et al., 2013). Such errors are typically examined within the domain of knowledge and expertise, such that people tend to be unaware of their low expertise (e.g., Dunning, Heath, & Suls, 2004). Similarly, there is evidence that intentions do not translate to behavior (Sheeran & Webb, 2016). Applied to smartphone cues, people are not good at estimating how often they get distracted, leading to the well-established discrepancy between perception of smartphone behavior and actual behavior (Boase & Ling, 2013; Scharrow, 2016; Vanden Abeele et al., 2013). In our case, it is plausible that participants perceived the smartphone cues during the experiment. Their own lay beliefs about the impact of smartphone cues may have led them to believe they were distracted. In reality, though, these cues might not have interfered with basic cognitive functions. In other words, the effect might have been strong enough to manifest in self-reports, but not in

behavior (Potter, 2011). However, this mechanism and other mechanisms discussed in Chapter 4 remain on the level of speculation until empirically tested.

INTERIM CONCLUSION 2

According to our results, smartphone cues do not appear problematic for executive control functions. Although participants feel that the presence of their smartphones or notifications distract them and put them in a state of heightened alertness, such cues did not impair performance. In two experiments, we found evidence that smartphone presence and smartphone notifications did not affect inhibition or working memory.

GOAL 3: ARE SMARTPHONE CUES REWARDING?

Even if smartphone cues do not impair executive control, why do people perceive them as distracting to such a degree that people experience alertness? One implicit, but untested assumption of many theories on smartphone behavior is that smartphone cues represent reward to users (Bayer et al., 2015; LaRose, 2015). People are social beings and have a strong need to connect to others (Baumeister & Leary, 1995; Deci & Ryan, 2000). Smartphones present people with a convenient way to gratify these social needs. According to those theories, people learn to associate the need gratification they obtain from their phones with smartphone cues (Le Pelley et al., 2016; Pool et al., 2016). Because people naturally seek out reward (Braver et al., 2014), their motivation for reward interacts with smartphone cues, such that smartphone cues become salient, attract attention, and trigger checking behavior (Anderson, 2016b, 2016a). Therefore, reward-associations could explain why people perceive smartphone cues as distracting and cause mental preoccupation with the online world (i.e., online vigilance).

In Chapter 6, we tested this proposition (Johannes, Dora, et al., 2019). Our results contest that smartphone-reward associations manifest behaviorally. Participants rated social smartphone apps (e.g., Facebook, Instagram) as much more rewarding than neutral smartphone apps (e.g., Calculator, Clock), particularly when these social smartphone apps had a notification sign. However, this difference in self-reported reward did not translate to behavior. Social smartphone apps high in reward did not attract more attention than apps low in reward or neutral apps. Moreover, half of the participants in the experiment were deprived of using their phones for one hour. We expected that deprivation would make smartphone apps even more rewarding, thus capturing attention to an even stronger degree. Instead, there was tentative evidence that deprived participants performed better than non-deprived participants.

Our results are in line with the findings of our previous studies. Yet again, we observed a strong discrepancy between people's perception and their behavior. We believe that this discrepancy is not due to low validity of what the self-reports of reward assessed. When instructing participants about what rewarding meant, we gave them several explanations

informed by the literature (e.g., feeling happy when using the app, having a strong urge to use the app). In addition, we applied several quality checks to make sure people read and processed these instructions. Consequently, we believe that the self-reports present an accurate measure of reward perception. On the side of the behavioral task, we relied on a procedure that has shown to be a valid measurement of reward associations (e.g., Anderson, 2016b). We therefore think that the discrepancy between ratings and attention responses is valid when assessing smartphone cues in the form of app icons. This discrepancy further supports our conclusion that smartphone cues and online vigilance are perceived as bothersome, but have little impact on performance or well-being.

The discrepancy between self-reported reward and the failure of that reward to capture attention has several potential explanations. Several theories propose that (rewarding) experiences are stored as complex situated conceptualizations (Barsalou, 2008; Papies & Barsalou, 2015). Smartphone cues, for example an app, will then be stored together with a participant's home screen, other apps surrounding that app icon, the mood the participant is in, as well as the environment. Encountering one cue, in our case the app icon, triggers people to simulate all other associated experiences, including reward. Such so-called "pattern completion inferences" (Papies & Barsalou, 2015, p. 37) work better with more cues. Consequently, isolating the app icon from its typical context might trigger feelings of reward only when participants have enough time to complete the pattern associated with the app and simulate the rewarding experience. During the survey, participants were free to take as much time as they needed. However, during the experiment, there was a time window, making it harder to fully process the app icon and its associated experiences. Alternatively, the discrepancy might yet again represent a self-evaluation error (Levine et al., 2013). People might hold lay theories about why they get distracted by apps, leading to high reported levels of reward. However, this self-report might not bear out for actual behavior.

INTERIM CONCLUSION 3

Our findings indicate that participants rate certain smartphone cues (i.e., apps) as rewarding. We did not find evidence that this reward association manifests in attention responses to the cues. Again, self-reported experience did not translate to performance. Whereas Chapters 4 and 5 showed a discrepancy between self-reported distraction and performance, Chapter 6 showed a discrepancy between reward ratings and the failure of those rewards to capture attention.

GOAL 4: CAN WE REDUCE PREFERENCES FOR SMARTPHONE CUES?

Despite a lack of evidence that smartphone cues impair performance, users still perceive them as rewarding. People perceive such appeal of smartphone cues as problematic

to resist, resulting in perceived distraction and a high temptation to procrastinate with media (Hofmann, Vohs, et al., 2012; Meier et al., 2016; Reinecke, Hartmann, & Eden, 2014; Reinecke & Hofmann, 2016). Although we could not show that these perceptions impair performance, perceived distraction likely still represents an unpleasant experience to users. Therefore, reducing the perceived appeal of smartphone cues represents a possible mechanism to tackle perceived distractions of smartphone cues.

In Chapter 7, we tested whether the so-called go/no-go training can reduce the appeal of smartphone cues. Specifically, we followed up on our findings in Chapter 6 where participants rated smartphone apps to be highly rewarding. Thus, we investigated whether the training can reduce explicit self-reported liking of smartphone apps. During the go/no-go training, participants responded to some apps (go trials), but not to others (no-go apps). We found that the training was effective. In Experiment 1, participants liked apps less if they did not respond to them, compared to responding or to an untrained baseline (i.e., apps not used in the training). In Experiment 2, we replicated this reduction in liking. In addition, we showed that liking played an important mediating role for choosing smartphone apps: If participants were presented with an app they had not responded to and an app they had responded to, their choice was partially driven by how much they liked each app. In other words, the training reduced how much participants liked certain apps; participants then relied on this liking to make a choice of what app to use.

Our results show that the go/no-go task presents a promising intervention tool for those people who wish to reduce the appeal of smartphone cues. The success of the training is particularly important given how famously difficult it is to change human behavior (Marteau et al., 2012; Sheeran & Webb, 2016), particularly media use. For instance, a recent review of multitasking interventions concluded that interventions to reduce media multitasking effects on cognition are rather limited in their effectiveness to change behavior or cognition (Parry & le Roux, 2019). Similarly, current interventions to reduce smartphone use mostly rely on a so-called nudging approach (Thaler & Sunstein, 2009). Rather than aiming to reduce preferences for certain apps that users perceive as distracting, nudging interventions implement design changes, such as reminders when people use an app longer than intended (Hiniker, Hong, Kohno, & Kientz, 2016). Although moderately effective in the immediate short-term, these nudges have not demonstrated persistent behavior change (Okeke, Sobolev, Dell, & Estrin, 2018). In contrast, there is strong evidence that directly targeting liking with the go/no-go task is effective to change behavior (Aulbach et al., 2019; Jones et al., 2016; Salomon et al., 2018; Schonberg et al., 2014; Turton et al., 2016). Consequently, the go/no-go task might also be an adequate intervention to reduce perceived distraction of smartphone cues.

Unlike in Chapter 6, the explicit appeal of apps translated to behavior in Chapter 7. Whereas high reward ratings did not result in attentional capture (Chapter 6), high appeal ratings influenced choices (Chapter 7). There are several possible reasons for this difference. First, the question of how rewarding an app is to participants (Chapter 6) might capture something different than the question of how attractive participants find

an app (Chapter 7). However, in our instructions for assessing reward, we explicitly told participants that rewarding means liking to use an app and that seeing the app makes them feel a strong need to open it. We believe that such features also apply to the appeal or attractiveness of an app. Second, in Chapter 7, participants rated many different apps, resulting in an idiosyncratic set of apps used in the training; in Chapter 6, they were presented with a preselection of different apps. Within a reward framework, the reward participants associate with apps depends on the participant's need and individual motivation (Botvinick & Braver, 2015). The preselection of apps in Chapter 6 might thus lead to a set that is not highly rewarding to participants. Both studies used a rating scale with identical range. Although the rating means across both studies were similar, there was more variation around the reward ratings in Chapter 6 compared to ratings in Chapter 7. This increased variation implies that having participants choose a tailored set of highly rewarding apps is more appropriate. In other words, this variation indicates that, although perceived as rewarding on the aggregate level, reward might not have been salient for many individual participants, obfuscating the effect on attentional capture.

INTERIM CONCLUSION 4

The results of our experiments suggest that the go/no-go training is effective in decreasing the self-reported appeal of smartphone cues. This reduction may be important; it has the promise to reduce the perceived distraction of smartphone cues and could even help influencing what smartphone cues people choose to engage with.

LIMITATIONS AND FUTURE DIRECTIONS

Our studies have several limitations. First, most of the studies we report here relied on an understanding that smartphone cues have the same meaning to all people. However, it is likely that people have different motivations and different types of usage that shape their reaction to smartphone cues. This limitation includes a lack of contextual and temporal dynamics: A notification might be highly distracting when bored, but not at all when fully engaged in a task (Eastwood et al., 2012; Kurzban, Duckworth, Kable, & Myers, 2013). In Chapter 7, smartphone apps were selected for the experimental procedure based on participants' explicit liking. Each participant received a tailored set of smartphone cues, which at least partially captures the differences between participants. Such procedures are simple to implement and present a more appropriate way to test the effect of smartphone cues compared to preselected sets of cues. In addition, we now have the technology to take factors of user motivation and temporal dynamics into account, such as building dynamic models combining logging data, activity trackers, and experience sampling over long periods of time.

Second, the complexity of media technology poses difficulties for experimental research. Many experimental researchers face the question whether high control

justifies highly artificial settings. In Chapters 4 and 5, we tried to strike a balance between control and artificiality by using participants' own phones. This enabled us to avoid the artificial situation of using a different phone. However, this decision required us to interact with the phones, likely raising suspicions among participants. In contrast, in Chapters 6 and 7, we required more control and used preselected app icons. This decision guaranteed standardization in stimuli across participants. However, it came at the cost of decontextualizing these apps. Consequently, it remains open to debate whether our insights about smartphone apps icons can be translated to regular behavior outside the lab. Yet again, answers to this problem of control versus ecological validity may lie in higher tailoring of stimulus sets – or to simply measure behavioral data. For example, the attention capture task of Chapter 6 could also be conducted on participants' own phones, relying on apps they have used a lot in the previous week. Alternatively, logging phone use over long periods of time combined with self-reports of distraction in the moment may yield more insightful results than lab studies. For instance, it is possible that the effect of smartphone cues on performance only occurs after prolonged exposure (Berkman, Hutcherson, Livingston, Kahn, & Inzlicht, 2017).

Third, it is unclear whether the negative effects of online vigilance are consequential for people's lives. In contrast to smartphone cues, we found some dimensions of online vigilance to relate negatively to well-being. However, these effects were small, in line with effects in the field in general (Rains et al., 2018) and within (social) media use and well-being specifically (Huang, 2017). The question is: Are those effects practically meaningful? Some authors propose that such effects are only of interest if people feel a difference in daily life. For example, Lakens, Scheel, and Isager (2018) list an example from medicine, where people noticed a difference in their lives only when the difference in their self-reports approached half a standard deviation. Accordingly, effects in Chapters 2 and 3 would need to be roughly three and half times larger before participants would truly notice a decline in their well-being. Translated to raw scores, participants would need to go from the midpoint of the online vigilance scale towards the maximum in order to feel a decline in their well-being. We need more research that moves beyond the simplistic distinction of significant versus not significant and inspects the actual impact of technology.

IMPLICATIONS AND FINAL THOUGHTS

Just like any previous technology that has had a drastic impact on society, smartphones are surrounded by controversy. The public and policy makers alike are concerned about possible negative effects of smartphones. Naturally, they turn to science for an answer. Remarkably, although it has been over ten years since the introduction of smartphones, we are still lacking coherent theoretical and empirical frameworks to provide such an answer. With this dissertation, we believe we made a first step towards such frameworks. We addressed the question of the effects of smartphone cues and online vigilance thoroughly

and comprehensively. First, we relied on a mix of cross-sectional, experimental, and field work. This approach can inform us of general trends, while testing whether such general trends also hold up in everyday life, complemented by tests of causality. Second, we distinguished between perception and objective behavior. This distinction proved to be necessary: Many theoretical assumptions about smartphone use almost exclusively concentrate on outcomes that manifest in behavior. However, our results show that perception and behavior are often disconnected. Theories about the effects of smartphone use can benefit from this insight. Incorporating perception into theory may allow us to draw a more nuanced, and possibly more accurate picture of the psychological consequences of smartphone cues and online vigilance. Third, we conducted methodologically rigorous, transparent empirical work that challenged claims made by previous work. As such, we demonstrate a need for the highest methodological and statistical rigor when investigating effects in our complex media environment.

So are concerns about this relatively new technology justified this time around? The research in this thesis showed that online vigilance has small effects on well-being when it represents absentmindedness. However, this negative effect does not seem important if this absentmindedness takes the form of positive thoughts. These findings have important theoretical and practical implications. Theoretically, our results suggest that the dimension of online vigilance that represents cognitive preoccupation can be understood as akin to other accounts describing the cognitive mechanisms behind absentmindedness. Both mind-wandering (Christoff et al., 2018; Smallwood & Schooler, 2015) and mindfulness (Brown & Ryan, 2003; Gu et al., 2015) share conceptual similarities with this salience dimension. These similarities become empirically evident in our studies, where thoughts about the online sphere behaved similarly to low mindfulness and high mind-wandering (Franklin et al., 2013; Gu et al., 2015; Killingsworth & Gilbert, 2010). Our results imply that dividing mental preoccupation into thoughts about the offline and online sphere might be artificial, because face-to-face and mediated communication are merging (Walther, 2017). Hence, suspecting thoughts about the online sphere to be more detrimental than thoughts about the offline sphere may put a misplaced premium on online communication (Ellis, 2019). This implication becomes particularly salient when considering that the other two dimensions, reactivity and monitoring, were not detrimental to well-being.

Is online vigilance problematic then? And what practical consequences can we draw based on the findings in this dissertation? To make it short: We do not believe smartphone users need to worry for now that their orientation toward online communication will have negative consequences for them. That is not to say online vigilance might not be problematic for some people – or for many more in the future. After all, science is cumulative and mostly interested in aggregate effects, not individuals. Our findings merely present a small step in understanding our connected lives. However, they show consistently that concerns are currently not warranted. Given that the valence of thoughts was by far the strongest factor for increasing well-being, it might be better to concentrate on maintaining one's relationship rather than worrying whether one is too preoccupied with online communication.

Even if online vigilance might not be problematic, the constant temptation of smartphone cues represents another concern for many actors in the public debate (e.g., Carr, 2011). Our findings suggest that such concerns, once again, may be premature. From a methodological perspective, our results demonstrate the need to conduct experiments to complement cross-sectional work. So far, cross-sectional work mostly paints a dim picture of the effect of smartphone cues on performance (Chein et al., 2017; Junco & Cotten, 2011). However, without experiments, it is unclear whether such relations are meaningful. Our results suggest that people, at least in the short-term, are not affected by smartphone cues in their performance. In fact, we found that perceived distraction is not necessarily related to performance (Levine et al., 2013). This discrepancy has both theoretical and practical implications. Theoretically, regarding people as passive agents who merely react to smartphone cues represents an old-fashioned understanding of media effects (Potter, 2011). Even though smartphone users perceive distraction, they seem to be quite good at ignoring it if necessary. This implies that we need to take motivational states and individual goals into account when investigating effects of mobile technology (Kurzman et al., 2013; Valkenburg & Peter, 2013). Practically, the discrepancy shows that we need to be careful when reacting to complaints about media technology. Complaints might not always be grounded in evidence.

Our results may also have important implications in the broader context of Social Science. We preregistered all hypotheses and tried to reduce researcher degrees of freedom to a minimum. Consequently, a large proportion of our research resulted in small or null effects. Such null effects are in line with research that employs preregistration and follows open science practices (Allen & Mehler, 2019). In a comparison, preregistered research reported less than half the effect size than that reported in non-preregistered research (Schäfer & Schwarz, 2019). Similarly, more recent work on the effects of media use on well-being finds negligible effects (Orben et al., 2019, 2019; Orben & Przybylski, 2019b) compared to previous work (Heffer et al., 2019). Our work follows a similar trajectory, often yielding small to negligible effects. Many researchers believe that small effects are just a natural consequence when investigating human behavior (Funder & Ozer, 2019). Just like in other fields of the Social Sciences, we are seeing the first failed replications appearing in the field of media effects (Keating & Totzkay, 2019; McEwan, Carpenter, & Westerman, 2018). We believe that such small effects will become more and more common in the field of media effects research with the adoption of open science practices.

Taking all limitations and implications into account, is there a need for policy makers to react? The scientific basis for making such policy recommendations does not appear robust. Currently, it would be premature to implement drastic measures, such as banning phones in many public spaces, at least not based on the findings of this dissertation. Smartphones have become a natural means of communication and people seem capable of ignoring smartphones' potential distractions. Consequently, it might be better to increase motivation for other tasks than to focus on common smartphone behavior. Although there have been concerns surrounding every new technology that rapidly gained

popularity, our results suggest that worries around smartphone cues and online vigilance are currently not justified. As such, our results are mostly inconsistent with previous, non-preregistered work. These inconsistencies in the literature warrant caution when making recommendations to policy makers. We need more high-powered, preregistered, open research before we can make such recommendations with a reasonable degree of certainty.

References

- Aarts, H., Custers, R., & Veltkamp, M. (2008). Goal priming and the affective-motivational route to nonconscious goal pursuit. *Social Cognition, 26*(5), 555–577. <https://doi.org/10.1521/soco.2008.26.5.555>
- Ajzen, I. (1991). The theory of planned behavior. *Organizational Behavior and Human Decision Processes, 50*(2), 179–211. [https://doi.org/10.1016/0749-5978\(91\)90020-T](https://doi.org/10.1016/0749-5978(91)90020-T)
- Albers, C. J., & Lakens, D. (2018). When power analyses based on pilot data are biased: Inaccurate effect size estimators and follow-up bias. *Journal of Experimental Social Psychology, 74*, 187–195. <https://doi.org/10.17605/OSF.IO/B7Z4Q>
- Allen, C., & Mehler, D. M. A. (2019). Open science challenges, benefits and tips in early career and beyond. *PLoS Biology, 17*(5), e3000246. <https://doi.org/10.1371/journal.pbio.3000246>
- Allom, V., Mullan, B., & Hagger, M. (2016). Does inhibitory control training improve health behaviour? A meta-analysis. *Health Psychology Review, 10*(2), 168–186. <https://doi.org/10.1080/17437199.2015.1051078>
- Alzahabi, R., & Becker, M. W. (2013). The association between media multitasking, task-switching, and dual-task performance. *Journal of Experimental Psychology-Human Perception and Performance, 39*(5), 1485–1495. <https://doi.org/10.1037/a0031208>
- Anderson, B. A. (2016a). Social reward shapes attentional biases. *Cognitive Neuroscience, 7*(1–4), 30–36. <https://doi.org/10.1080/17588928.2015.1047823>
- Anderson, B. A. (2016b). The attention habit: How reward learning shapes attentional selection. *Annals of the New York Academy of Sciences, 1369*(1), 24–39. <https://doi.org/10.1111/nyas.12957>
- Anderson, B. A., Laurent, P. a, & Yantis, S. (2011a). Value-driven attentional capture. *Proceedings of the National Academy of Sciences, 108*(25), 10367–10371. <https://doi.org/10.1073/pnas.1104047108>
- Anderson, B. A., Laurent, P. A., & Yantis, S. (2011b). Learned value magnifies salience-based attentional capture. *PLoS ONE, 6*(11). <https://doi.org/10.1371/journal.pone.0027926>
- 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), 1–9. <https://doi.org/10.1371/journal.pone.0139004>
- Antheunis, M. L., Vanden Abeele, M. M. P., & Kanters, S. (2015). The impact of Facebook use on micro-level social capital: A synthesis. *Societies, 5*(2), 399–419. <https://doi.org/10.3390/soc5020399>
- Aulbach, M. B., Knittle, K., & Haukkala, A. (2019). Implicit process interventions in eating behaviour: A meta-analysis examining mediators and moderators. *Health Psychology Review, 1*–30. <https://doi.org/10.1080/17437199.2019.1571933>
- Barr, D. J., Levy, R., Scheepers, C., & Tily, H. J. (2013). Random effects structure for confirmatory hypothesis testing: Keep it maximal. *Journal of Memory and Language, 68*(3), 255–278. <https://doi.org/10.1016/j.jml.2012.11.001>
- Barsalou, L. W. (2008). Grounded cognition. *Annual Review of Psychology, 59*(1), 617–645. <https://doi.org/10.1146/annurev.psych.59.103006.093639>

- Barsalou, L. W. (2009). Simulation, situated conceptualization, and prediction. *Philosophical Transactions of the Royal Society B: Biological Sciences*, 364(1521), 1281–1289. <https://doi.org/10.1098/rstb.2008.0319>
- Barton, K. (2018). *MuMIn: Multi-Model Inference*. Retrieved from <https://CRAN.R-project.org/package=MuMIn>
- Bates, D., Maechler, 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>
- Bauer, A. A., Loy, L. S., Masur, P. K., & Schneider, F. M. (2017). Mindful instant messaging. *Journal of Media Psychology*, 29(3), 159–165. <https://doi.org/10.1027/1864-1105/a000225>
- Baumeister, R. F., & Leary, M. R. (1995). The need to belong: Desire for interpersonal attachments as a fundamental human motivation. *Psychological Bulletin*, 117(3), 497–529. <https://doi.org/10.1037/0033-2909.117.3.497>
- Baumgartner, S. E., van der Schuur, W. A., Lemmens, J. S., & te Poel, F. (2017). The relationship between media multitasking and attention problems in adolescents: Results of two longitudinal studies. *Human Communication Research*, 1–27. <https://doi.org/10.1111/hcre.12111>
- Bayer, J. B., & Campbell, S. W. (2012). Texting while driving on automatic: Considering the frequency-independent side of habit. *Computers in Human Behavior*, 28(6), 2083–2090. <https://doi.org/10.1016/j.chb.2012.06.012>
- Bayer, J. B., Campbell, S. W., & Ling, R. (2015). 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., Dal Cin, S., Campbell, S. W., & Panek, E. (2016). Consciousness and self-regulation in mobile communication. *Human Communication Research*, 42, 71–97. <https://doi.org/10.1111/hcre.12067>
- Bayer, J. B., Ellison, N., Schoenebeck, S., Brady, E., & Falk, E. B. (2018). Facebook in context(s): Measuring emotional responses across time and space. *New Media and Society*, 20(3), 1047–1067. <https://doi.org/10.1177/1461444816681522>
- Bayer, J. B., & LaRose, R. (2018). Technology habits: Progress, problems, and prospects. In B. Verplanken (Ed.), *The Psychology of Habit* (pp. 111–130). https://doi.org/10.1007/978-3-319-97529-0_7
- Beauchamp, K. G., Kahn, L. E., & Berkman, E. T. (2016). Does inhibitory control training transfer? Behavioral and neural effects on an untrained emotion regulation task. *Social Cognitive and Affective Neuroscience*, 11(9), 1374–1382. <https://doi.org/10.1093/scan/nsw061>
- Benjamin, D. J., Berger, J. O., Johannesson, M., Nosek, B. A., Wagenmakers, E.-J., Berk, R., ... Johnson, V. E. (2018). Redefine statistical significance. *Nature Human Behaviour*, 2(1), 6–10. <https://doi.org/10.1038/s41562-017-0189-z>
- Berkman, E. T., Hutcherson, C. A., Livingston, J. L., Kahn, L. E., & Inzlicht, M. (2017). Self-control as value-based choice. *Current Directions in Psychological Science*, 26(5), 422–428.

- Best, J. R., Miller, P. H., & Naglieri, J. A. (2011). Relations between executive function and academic achievement from ages 5 to 17 in a large, representative national sample. *Learning and Individual Differences, 21*(4), 327–336. <https://doi.org/10.1016/j.lindif.2011.01.007>
- Best, M., Lawrence, N. S., Logan, G. D., McLaren, I. P. L., & Verbruggen, F. (2016). Should I stop or should I go? The role of associations and expectancies. *Journal of Experimental Psychology: Human Perception and Performance, 42*(1), 115–137. <https://doi.org/10.1037/xhp0000116>
- Best, M., & Papies, E. K. (2017). Right here, right now: Situated interventions to change consumer habits. *Journal of the Association for Consumer Research, 2*(3), 333–358. <https://doi.org/10.1086/695443>
- Birnholtz, J., Davison, J., & Li, A. (2017). Attending to attention: How do people attract, manage, and negotiate attention using mobile devices? *Mobile Media and Communication, 5*(3), 256–274. <https://doi.org/10.1177/2050157917714504>
- bitkom. (2015). Zahl der verschickten SMS sinkt um 40 Prozent [Number of sent SMS decreases by 40%]. Retrieved from <https://www.bitkom.org/Presse/Presseinformation/Zahl-der-verschickten-SMS-sinkt-um-40-Prozent.html>
- Blask, K., Frings, C., & Walther, E. (2016). Doing is for feeling. *Journal of Experimental Psychology: General, 145*(10), 1263–1268. <https://doi.org/10.1037/xge0000211>
- Boase, J., & Ling, R. (2013). Measuring mobile phone use: Self-report versus log data. *Journal of Computer-Mediated Communication, 18*(4), 508–519.
- Bohlmeijer, E., ten Klooster, P. M., Fledderus, M., Veehof, M., & Baer, R. (2011). Psychometric properties of the five facet mindfulness questionnaire in depressed adults and development of a short form. *Assessment, 18*(3), 308–320. <https://doi.org/10.1177/1073191111408231>
- Botvinick, M., & Braver, T. (2015). Motivation and cognitive control: From behavior to neural mechanism. *Annual Review of Psychology, 66*(1), 83–113. <https://doi.org/10.1146/annurev-psych-010814-015044>
- Bowman, L. L., Levine, L. E., Waite, B. M., & Gendron, M. (2010). Can students really multitask? An experimental study of instant messaging while reading. *Computers & Education, 54*(4), 927–931. <https://doi.org/10.1016/j.compedu.2009.09.024>
- Braver, T. S., Krug, M. K., Chiew, K. S., Kool, W., Westbrook, J. A., Clement, N. J., ... Somerville, L. H. (2014). Mechanisms of motivation-cognition interaction: Challenges and opportunities. *Cognitive, Affective & Behavioral Neuroscience, 14*(2), 443–472. <https://doi.org/10.3758/s13415-014-0300-0>
- Brown, K. W., & Ryan, R. M. (2003). The benefits of being present: Mindfulness and its role in psychological well-being. *Journal of Personality and Social Psychology, 84*(4), 822–848. <https://doi.org/10.1037/0022-3514.84.4.822>
- Brown, K. W., Ryan, R. M., & Creswell, J. D. (2007). Mindfulness: Theoretical foundations and evidence for its salutary effects. *Psychological Inquiry, 18*(March), 211–237. <https://doi.org/10.1080/10478400701598298>
- Brysbaert, M., & Stevens, M. A. (2018). Power analysis and effect size in mixed effects models: A tutorial. *Journal of Cognition, 1*(1), 1–20. <https://doi.org/10.5334/joc.10>

- Burke, M., & Kraut, R. E. (2016). The relationship between Facebook use and well-being depends on communication type and tie strength. *Journal of Computer-Mediated Communication*, 21(4), 265–281. <https://doi.org/10.1111/jcc4.12162>
- Bürkner, P.-C. (2017). brms: An R package for Bayesian multilevel models using Stan. *Journal of Statistical Software*, 80(1). <https://doi.org/10.18637/jss.v080.i01>
- Button, K. S., Ioannidis, J. P. a, Mokrysz, C., Nosek, B. a, Flint, J., Robinson, E. S. J., & Munafò, M. R. (2013). Power failure: Why small sample size undermines the reliability of neuroscience. *Nature Reviews. Neuroscience*, 14(5), 365–376. <https://doi.org/10.1038/nrn3475>
- Calderwood, C., Ackerman, P. L., & Conklin, E. M. (2014). What else do college students “do” while studying? An investigation of multitasking. *Computers and Education*, 75, 19–29. <https://doi.org/10.1016/j.compedu.2014.02.004>
- Camerer, C. F., Dreber, A., Holzmeister, F., Ho, T.-H., Huber, J., Johannesson, M., ... Wu, H. (2018). Evaluating the replicability of social science experiments in Nature and Science between 2010 and 2015. *Nature Human Behaviour*. <https://doi.org/10.1038/s41562-018-0399-z>
- Carbine, K. A., & Larson, M. J. (2019). Quantifying the presence of evidential value and selective reporting in food-related inhibitory control training: A p-curve analysis. *Health Psychology Review*, 1–26. <https://doi.org/10.1080/17437199.2019.1622144>
- Carolus, A., Binder, J. F., Muench, R., Schmidt, C., Schneider, F., & Buglass, S. L. (2018). Smartphones as digital companions: Characterizing the relationship between users and their phones. *New Media & Society*, 1–25. <https://doi.org/10.1177/1461444818817074>
- Carr, N. (2011). *The shallows—How the internet is changing the way we think, read and remember*. London: Atlantic Books.
- CBS. (2018). Internet; toegan, gebruik en faciliteiten [Internet; access, use and devices]. Retrieved from <https://opendata.cbs.nl/statline/#/CBS/nl/dataset/83429NED/table?ts=1531831858987>
- Chambers, C. D., Dienes, Z., McIntosh, R. D., Rotshtein, P., & Willmes, K. (2015). Registered Reports: Realigning incentives in scientific publishing. *Cortex*, 66, A1–A2. <https://doi.org/10.1016/j.cortex.2015.03.022>
- Cheever, N. A., Rosen, L. D., Carrier, L. M., & Chavez, A. (2014). Out of sight is not out of mind: The impact of restricting wireless mobile device use on anxiety levels among low, moderate and high users. *Computers in Human Behavior*, 37, 290–297. <https://doi.org/10.1016/j.chb.2014.05.002>
- Chein, J. M., Wilmer, H. H., & Sherman, L. E. (2017). Smartphones and cognition: A review of research exploring the links between mobile technology habits and cognitive functioning. *Frontiers in Psychology*, 8, 1–16. <https://doi.org/10.3389/fpsyg.2017.00605>
- Chelazzi, L., Perlato, A., Santandrea, E., & Della Libera, C. (2013). Rewards teach visual selective attention. *Vision Research*, 85, 58–62. <https://doi.org/10.1016/j.visres.2012.12.005>
- Chen, Q., & Yan, Z. (2016). Does multitasking with mobile phones affect learning? A review. *Computers in Human Behavior*, 54, 34–42. <https://doi.org/10.1016/j.chb.2015.07.047>

- Chen, Z., Holland, R. W., Quandt, J., Dijksterhuis, A., & Veling, H. (2019). When mere action versus inaction leads to robust preference change. *Journal of Personality and Social Psychology*. <https://doi.org/0.1037/pspa0000158>
- Chen, Z., Veling, H., Dijksterhuis, A., & Holland, R. W. (2016). How does not responding to appetitive stimuli cause devaluation: Evaluative conditioning or response inhibition? *Journal of Experimental Psychology: General*, *145*(12), 1687–1701. <https://doi.org/10.1037/xge0000236>
- Christoff, K., Mills, C., Andrews-Hanna, J. R., Irving, Z. C., Thompson, E., Fox, K. C. R., & Kam, J. W. Y. (2018). Mind-wandering as a scientific concept: Cutting through the definitional haze. *Trends in Cognitive Sciences*, *22*(11), 967–959. <https://doi.org/10.1016/j.tics.2018.07.004>
- Chun, H., Lee, H., & Kim, D. (2012). The integrated model of smartphone adoption: Hedonic and utilitarian value perceptions of smartphones among Korean college students. *Cyberpsychology, Behavior, and Social Networking*, *15*(9), 473–479. <https://doi.org/10.1089/cyber.2012.0140>
- Clancy, E. M., Fiacconi, C. M., & Fenske, M. J. (2019). Response inhibition immediately elicits negative affect and devalues associated stimuli: Evidence from facial electromyography. In *Progress in Brain Research*. <https://doi.org/10.1016/bs.pbr.2019.03.019>
- Clayton, R. B., Leshner, G., & Almond, A. (2015). The extended iSelf: The impact of iPhone separation on cognition, emotion, and physiology. *Journal of Computer-Mediated Communication*, *20*, 119–135. <https://doi.org/10.1111/jcc4.12109>
- Cserjési, R., Luminet, O., Poncelet, A.-S., & Lénárd, L. (2009). Altered executive function in obesity. Exploration of the role of affective states on cognitive abilities. *Appetite*, *52*(2), 535–539. <https://doi.org/10.1016/j.appet.2009.01.003>
- Csikszentmihalyi, M., & Larson, R. (2014). Validity and Reliability of the Experience-Sampling Method. In M. Csikszentmihalyi (Ed.), *Flow and the Foundations of Positive Psychology: The Collected Works of Mihaly Csikszentmihalyi* (pp. 35–54). https://doi.org/10.1007/978-94-017-9088-8_3
- Cumiskey, K. M., & Ling, R. (2015). The social psychology of mobile communication. In S. S. Sundar (Ed.), *The handbook of the psychology of communication technology* (pp. 228–246). Chichester, UK: John Wiley & Sons, Inc.
- Cutino, C. M., & Nees, M. A. (2016). Restricting mobile phone access during homework increases attainment of study goals. *Mobile Media & Communication*. <https://doi.org/10.1177/2050157916664558>
- De Groot, A. D. (2014). The meaning of “significance” for different types of research [translated and annotated by Eric-Jan Wagenmakers, Denny Borsboom, Josine Verhagen, Rogier Kievit, Marjan Bakker, Angelique Cramer, Dora Matzke, Don Mellenbergh, and Han L. J. van der Maas]. *Acta Psychologica*, *148*, 188–194. <https://doi.org/10.1016/j.actpsy.2014.02.001>
- De Houwer, J. (2007). A conceptual and theoretical analysis of evaluative conditioning. *The Spanish Journal of Psychology*, *10*(02), 230–241. <https://doi.org/10.1017/S1138741600006491>
- Deci, E. L., & Ryan, R. M. (2000). The “what” and “why” of goal pursuits: Human needs and the self-determination of behavior. *Psychological Inquiry*, *11*(4), 227–268. https://doi.org/10.1207/S15327965PLI1104_01

- Diamond, A. (2014). Executive functions. *Annual Review of Clinical Psychology*, *64*, 135–168. <https://doi.org/10.1146/annurev-psych-113011-143750>
- Diener, E., Emmons, R. A., Larsen, R. J., & Griffin, S. (1985). The satisfaction with life scale. *Journal of Personality Assessment*, *49*(1), 71–75.
- Diener, E., Lucas, R. E., & Oishi, S. (2018). Advances and open questions in the science of subjective well-being. *Collabra: Psychology*, *4*(1), 15. <https://doi.org/10.1525/collabra.115>
- Diener, E., Wirtz, D., Tov, W., Kim-Prieto, C., Choi, D., Oishi, S., & Biswas-Diener, R. (2010). New well-being measures: Short scales to assess flourishing and positive and negative feelings. *Social Indicators Research*, *97*(2), 143–156. <https://doi.org/10.1007/s11205-009-9493-y>
- Dietz, S., & Henrich, C. (2014). Texting as a distraction to learning in college students. *Computers in Human Behavior*, *36*, 163–167. <https://doi.org/10.1016/j.chb.2014.03.045>
- Domahidi, E. (2018). The associations between online media use and users' perceived social resources: A meta-analysis. *Journal of Computer-Mediated Communication*, *23*(4), 181–200. <https://doi.org/10.1093/jcmc/zmy007>
- Driscoll, R. L., de Launay, K. Q., & Fenske, M. J. (2018). Less approach, more avoidance: Response inhibition has motivational consequences for sexual stimuli that reflect changes in affective value not a lingering global brake on behavior. *Psychonomic Bulletin & Review*, *25*(1), 463–471. <https://doi.org/10.3758/s13423-017-1291-y>
- Du, J., Kerkhof, P., & van Koningsbruggen, G. M. (2019). Predictors of social media self-control failure: Immediate gratifications, habitual checking, ubiquity, and notifications. *Cyberpsychology, Behavior, and Social Networking*, *22*(7), 477–485. <https://doi.org/10.1089/cyber.2018.0730>
- Dunning, D., Heath, C., & Suls, J. M. (2004). Flawed self-assessment: Implications for health, education, and the workplace. *Psychological Science in the Public Interest*, *5*(3), 69–106. <https://doi.org/10.1111/j.1529-1006.2004.00018.x>
- Eastwood, J., Frischen, A., Fenske, M. J., & Smilek, D. (2012). The unengaged mind: Defining boredom in terms of attention. *Perspectives on Psychological Science*, *7*(5), 482–495. <https://doi.org/10.1177/1745691612456044>
- Eder, A. B., Krishna, A., & Van Dessel, P. (2019). Operant evaluative conditioning. *Journal of Experimental Psychology: Animal Learning and Cognition*, *45*(1), 102–110. <https://doi.org/10.1037/xan0000189>
- Elhai, J. D., & Contractor, A. A. (2018). Examining latent classes of smartphone users: Relations with psychopathology and problematic smartphone use. *Computers in Human Behavior*, *82*, 159–166. <https://doi.org/10.1016/j.chb.2018.01.010>
- Elhai, J. D., Levine, J. C., O'Brien, K. D., & Armour, C. (2018). Distress tolerance and mindfulness mediate relations between depression and anxiety sensitivity with problematic smartphone use. *Computers in Human Behavior*, *84*, 477–484. <https://doi.org/10.1016/j.chb.2018.03.026>
- Ellis, D. A. (2019). Are smartphones really that bad? Improving the psychological measurement of technology-related behaviors. *Computers in Human Behavior*. <https://doi.org/10.1016/j.chb.2019.03.006>

- Ellis, D. A., Davidson, B. I., Shaw, H., & Geyer, K. (2019). Do smartphone usage scales predict behavior? *International Journal of Human-Computer Studies*, *130*, 86–92. <https://doi.org/10.1016/j.ijhcs.2019.05.004>
- Enge, S., Behnke, A., Fleischhauer, M., Küttler, L., Kliegel, M., & Strobel, A. (2014). No evidence for true training and transfer effects after inhibitory control training in young healthy adults. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *40*(4), 987–1001. <https://doi.org/10.1037/a0036165>
- Epstein, L. H., Truesdale, R., Wojcik, A., Paluch, R. A., & Raynor, H. A. (2003). Effects of deprivation on hedonics and reinforcing value of food. *Physiology & Behavior*, *78*, 221–227. [https://doi.org/10.1016/S0031-9384\(02\)00978-2](https://doi.org/10.1016/S0031-9384(02)00978-2)
- Failing, M. F., & Theeuwes, J. (2017). Selection history: How reward modulates selectivity of visual attention. *Psychonomic Bulletin and Review*, *25*(2), 514–538. <https://doi.org/10.3758/s13423-017-1380-y>
- Forman, E. M., Manasse, S. M., Dallal, D. H., Crochiere, Rebecca, J., Loyka, C. M., Butryn, M. L., ... Houben, K. (2019). Computerized neurocognitive training for improving dietary health and facilitating weight loss. *Journal of Behavioral Medicine*. <https://doi.org/10.1007/s10865-019-00024-5>
- Franklin, M. S., Mrazek, M. D., Anderson, C. L., Smallwood, J., Kingstone, A., & Schooler, J. (2013). The silver lining of a mind in the clouds: Interesting musings are associated with positive mood while mind-wandering. *Frontiers in Psychology*, *4*, 583.
- Friedman, N. P., Miyake, A., Young, S. E., Defries, J. C., Corley, R. P., & Hewitt, J. K. (2008). Individual differences in executive functions are almost entirely genetic in origin. *Journal of Experimental Psychology: General*, *137*(2), 201–225. <https://doi.org/10.1037/0096-3445.137.2.201>
- Friese, M., & Hofmann, W. (2016). State mindfulness, self-regulation, and emotional experience in everyday life. *Motivation Science*, *2*(1), 1–14. <https://doi.org/10.1037/mot0000027>
- Funder, D. C., & Ozer, D. J. (2019). Evaluating effect size in psychological research: Sense and nonsense. *Advances in Methods and Practices in Psychological Science*, *2*(2), 156–168. <https://doi.org/10.1177/2515245919847202>
- Gelman, A., & Loken, E. (2013). The garden of forking paths: Why multiple comparisons can be a problem, even when there is no “fishing expedition” or “p-hacking” and the research hypothesis was posited ahead of time. *Department of Statistics, Columbia University*.
- Giambra, L. M. (1993). The influence of aging on spontaneous shifts of attention from external stimuli to the contents of consciousness. *Experimental Gerontology*, *28*(4–5), 485–492. [https://doi.org/10.1016/0531-5565\(93\)90073-M](https://doi.org/10.1016/0531-5565(93)90073-M)
- Green, P., & Macleod, C. J. (2016). simr: An R package for power analysis of generalized linear mixed models by simulation. *Methods in Ecology and Evolution*, *7*(4), 493–498. <https://doi.org/10.1111/2041-210X.12504>
- Gu, J., Strauss, C., Bond, R., & Cavanagh, K. (2015). How do mindfulness-based cognitive therapy and mindfulness-based stress reduction improve mental health and wellbeing? A systematic review and meta-analysis of mediation studies. *Clinical Psychology Review*, *37*, 1–12. <https://doi.org/10.1016/j.cpr.2015.01.006>

- Guitart-Masip, M., Duzel, E., Dolan, R., & Dayan, P. (2014). Action versus valence in decision making. *Trends in Cognitive Sciences, 18*(4), 194–202. <https://doi.org/10.1016/j.tics.2014.01.003>
- Guitart-Masip, M., Huys, Q. J. M., Fuentemilla, L., Dayan, P., Duzel, E., & Dolan, R. J. (2012). Go and no-go learning in reward and punishment: Interactions between affect and effect. *NeuroImage, 62*(1), 154–166. <https://doi.org/10.1016/j.neuroimage.2012.04.024>
- Hadlington, L. J. (2015). Cognitive failures in daily life: Exploring the link with Internet addiction and problematic mobile phone use. *Computers in Human Behavior, 51*, 75–81. <https://doi.org/10.1016/j.chb.2015.04.036>
- Halekoh, U., & Højsgaard, S. (2014). A Kenward-Roger approximation and parametric bootstrap methods for tests in linear mixed models – The R package pbkrtest. *Journal of Statistical Software, 59*(9), 1–30.
- 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. A., & Baym, N. K. (2012). Calling and texting (too much): Mobile maintenance expectations, (over)dependence, entrapment, and friendship satisfaction. *New Media & Society, 14*(2), 316–331. <https://doi.org/10.1177/1461444811415047>
- Hamaker, E. L., Kuiper, R. M., & Grasman, R. P. P. P. (2015). A critique of the cross-lagged panel model. *Psychological Methods, 20*(1), 102–116. <https://doi.org/10.1037/a0038889>
- Hartanto, A., & Yang, H. (2016). Is the smartphone a smart choice? The effect of smartphone separation on executive functions. *Computers in Human Behavior, 64*, 329–336. <https://doi.org/10.1016/j.chb.2016.07.002>
- Hayes, A. F., & Scharkow, M. (2013). The relative trustworthiness of inferential tests of the indirect effect in statistical mediation analysis: Does method really matter? *Psychological Science, 24*(10), 1918–1927. <https://doi.org/10.1177/0956797613480187>
- Heffer, T., Good, M., Daly, O., MacDonell, E., & Willoughby, T. (2019). The longitudinal association between social-media use and depressive symptoms among adolescents and young adults: An empirical reply to Twenge et al. (2018). *Clinical Psychological Science, 2167702618812727*. <https://doi.org/10.1177/2167702618812727>
- Helliwell, J. F., & Wang, S. (2014). Weekends and subjective well-being. *Social Indicators Research, 116*(2), 389–407. <https://doi.org/10.1007/s11205-013-0306-y>
- Hickey, C., Chelazzi, L., & Theeuwes, J. (2010). Reward changes salience in human vision via the Anterior Cingulate. *Journal of Neuroscience, 30*(33), 11096–11103. <https://doi.org/10.1523/JNEUROSCI.1026-10.2010>
- Hickey, C., Kaiser, D., & Peelen, M. V. (2015). Reward guides attention to object categories in real-world scenes. *Journal of Experimental Psychology: General, 144*(2), 264–273. <https://doi.org/10.1037/a0038627>
- Hiniker, A., Hong, S. (Ray), Kohno, T., & Kientz, J. A. (2016). MyTime: Designing and evaluating an intervention for smartphone non-use. *Proceedings of the 2016 CHI Conference on Human Factors in Computing Systems, 4746–4757*. <https://doi.org/10.1145/2858036.2858403>

- Hintze, D., Hintze, P., Findling, R. D., & Mayrhofer, R. (2017). A large-scale, long-term analysis of mobile device usage characteristics. *Proceedings of the ACM on Interactive, Mobile, Wearable and Ubiquitous Technologies*, *1*(2), 1–21. <https://doi.org/10.1145/3090078>
- Hoffner, C. a., & Lee, S. (2015). Mobile phone use, emotion regulation, and well-being. *Cyberpsychology, Behavior, and Social Networking*, *18*(7), 411–416. <https://doi.org/10.1089/cyber.2014.0487>
- Hofmann, W., Schmeichel, B. J., & Baddeley, A. D. (2012). Executive functions and self-regulation. *Trends in Cognitive Sciences*, *16*(3), 174–180. <https://doi.org/10.1016/j.tics.2012.01.006>
- Hofmann, W., Vohs, K. D., & Baumeister, R. F. (2012). What people desire, feel conflicted about, and try to resist in everyday life. *Psychological Science*, *23*(6), 582–588. <https://doi.org/10.1177/0956797612437426>
- Hollis, R. B., & Was, C. A. (2016). Mind wandering, control failures, and social media distractions in online learning. *Learning and Instruction*, *42*, 104–112. <https://doi.org/10.1016/j.learninstruc.2016.01.007>
- Hommel, B. (2004). Event files: Feature binding in and across perception and action. *Trends in Cognitive Sciences*, *8*(11), 494–500. <https://doi.org/10.1016/j.tics.2004.08.007>
- Houben, K., Havermans, R. C., Nederkoorn, C., & Jansen, A. (2012). Beer à no-go: Learning to stop responding to alcohol cues reduces alcohol intake via reduced affective associations rather than increased response inhibition. *Addiction*, *107*(7), 1280–1287. <https://doi.org/10.1111/j.1360-0443.2012.03827.x>
- Houben, K., & Jansen, A. (2011). Training inhibitory control. A recipe for resisting sweet temptations. *Appetite*, *56*(2), 345–349. <https://doi.org/10.1016/j.appet.2010.12.017>
- Huang, C. (2017). Time spent on social network sites and psychological well-being: A meta-analysis. *Cyberpsychology, Behavior, and Social Networking*, *20*(6), 346–354. <https://doi.org/10.1089/cyber.2016.0758>
- Hughes, S., De Houwer, J., & Perugini, M. (2016). Expanding the boundaries of evaluative learning research: How intersecting regularities shape our likes and dislikes. *Journal of Experimental Psychology: General*, *145*(6), 731–754. <https://doi.org/10.1037/xge0000100>
- Inzlicht, M., & Berkman, E. (2015). Six questions for the resource model of control (and some answers). *Social and Personality Psychology Compass*, *9*(10), 511–524. <https://doi.org/10.1111/spc3.12200>
- Ishii, K., Rife, T. S., & Kagawa, N. (2017). Technology-driven gratifications sought through text-messaging among college students in the U.S. and Japan. *Computers in Human Behavior*, *69*, 396–404. <https://doi.org/10.1016/j.chb.2016.12.022>
- Jain, S., Shapiro, S. L., Swanick, S., Roesch, S. C., Mills, P. J., Bell, I., & Schwartz, G. E. R. (2007). A randomized controlled trial of mindfulness meditation versus relaxation training: Effects on distress, positive states of mind, rumination, and distraction. *Annals of Behavioral Medicine*, *33*(1), 11–21. https://doi.org/10.1207/s15324796abm3301_2
- JASP Team. (2017). *JASP (Version 0.8.1.1)[Computer software]*.
- Jeong, S.-H., & Hwang, Y. (2016). Media multitasking effects on cognitive vs. Attitudinal outcomes: A meta-analysis. *Human Communication Research*, 1–20. <https://doi.org/10.1111/hcre.12089>

- Johannes, N., Dora, J., & Rusz, D. (2019). Social smartphone apps do not capture attention despite their perceived high reward value. *Collabra: Psychology*, 5(1), 14. <https://doi.org/10.1525/collabra.207>
- Johannes, N., Veling, H., & Buijzen, M. (2019). *No evidence that smartphone notifications lead to goal-neglect* [Preprint]. <https://doi.org/10.31234/osf.io/5me97>
- 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. (in press). Hard to resist? The effect of smartphone visibility and notifications on response inhibition. *Journal of Media Psychology, Advance Online Publication*. <https://doi.org/10.1027/1864-1105/a000248>
- John, L. K., Loewenstein, G., & Prelec, D. (2012). Measuring the prevalence of questionable research practices with incentives for truth telling. *Psychological Science*, 23(5), 524–532. <https://doi.org/10.1177/0956797611430953>
- Jones, A., Di Lemma, L. C. G., Robinson, E., Christiansen, P., Nolan, S., Tudur-Smith, C., & Field, M. (2016). Inhibitory control training for appetitive behaviour change: A meta-analytic investigation of mechanisms of action and moderators of effectiveness. *Appetite*, 97, 16–28. <https://doi.org/10.1016/j.appet.2015.11.013>
- Junco, R., & Cotten, S. R. (2011). Perceived academic effects of instant messaging use. *Computers & Education*, 56(2), 370–378. <https://doi.org/10.1016/j.compedu.2010.08.020>
- Jung, E. H., & Sundar, S. S. (2018). Status update: Gratifications derived from Facebook affordances by older adults. *New Media and Society*, 20(11), 4135–4154. <https://doi.org/10.1177/1461444818768090>
- Jurado, M. B., & Rosselli, M. (2007). The elusive nature of executive functions: A review of our current understanding. *Neuropsychology Review*, 17(3), 213–233. <https://doi.org/10.1007/s11065-007-9040-z>
- Kane, M. J., & Engle, R. W. (2003). Working-memory capacity and the control of attention: The contributions of goal neglect, response competition, and task set to Stroop interference. *Journal of Experimental Psychology: General*, 132(1), 47–70. <https://doi.org/10.1037/0096-3445.132.1.47>
- Kanjo, E., Kuss, D. J., & Ang, C. S. (2017). NotiMind: Utilizing responses to smartphone notifications as affective sensors. *IEEE Access*, 5, 22023–22035. <https://doi.org/10.1109/ACCESS.2017.2755661>
- Karapanos, E., Teixeira, P., & Gouveia, R. (2016). Need fulfillment and experiences on social media: A case on Facebook and WhatsApp. *Computers in Human Behavior*, 55, 888–897. <https://doi.org/10.1016/j.chb.2015.10.015>
- Kardefelt-Winther, D., Heeren, A., Schimmenti, A., van Rooij, A., Maurage, P., Carras, M., ... Billieux, J. (2017). How can we conceptualize behavioural addiction without pathologizing common behaviours? *Addiction*, 112(10), 1709–1715. <https://doi.org/10.1111/add.13763>

- Kardos, P., Unoka, Z., Pléh, C., & Soltész, P. (2018). Your mobile phone indeed means your social network: Priming mobile phone activates relationship related concepts. *Computers in Human Behavior*, *88*, 84–88. <https://doi.org/10.1016/j.chb.2018.06.027>
- Katevas, K., Arapakis, I., & Pielot, M. (2018). Typical phone use habits: Intense use does not predict negative well-being. *Proceedings of the 20th International Conference on Human-Computer Interaction with Mobile Devices and Services - MobileHCI '18*, 1–13. <https://doi.org/10.1145/3229434.3229441>
- Keating, D. M., & Totzkay, D. (2019). We do publish (conceptual) replications (sometimes): Publication trends in communication science, 2007–2016. *Annals of the International Communication Association*, *43*(3), 225–239. <https://doi.org/10.1080/23808985.2019.1632218>
- Keng, S. L., Smoski, M. J., & Robins, C. J. (2011). Effects of mindfulness on psychological health: A review of empirical studies. *Clinical Psychology Review*, *31*(6), 1041–1056. <https://doi.org/10.1016/j.cpr.2011.04.006>
- Kenny, D. A., Kashy, D. A., & Bolger, N. (1998). Data analysis in social psychology. In D. T. Gilbert, S. T. Fiske, & G. Lindzey (Eds.), *The handbook of social psychology* (4th ed., Vol. 1, pp. 233–265). New York: Oxford University Press.
- Killingsworth, M. A., & Gilbert, D. T. (2010). A wandering mind is an unhappy mind. *Science*, *330*(6006), 932–932. <https://doi.org/10.1126/science.1192439>
- Klimmt, C., Hefner, D., Reinecke, L., Rieger, D., & Vorderer, P. (2018). The permanently online and permanently connected mind. Mapping the cognitive structures behind mobile Internet use. In P. Vorderer, D. Hefner, L. Reinecke, & C. Klimmt (Eds.), *Permanently online, permanently connected. Living and communication in a POPC world* (pp. 18–28). New York: Routledge.
- Koch, S., Holland, R. W., & van Knippenberg, A. (2008). Regulating cognitive control through approach-avoidance motor actions. *Cognition*, *109*(1), 133–142. <https://doi.org/10.1016/j.cognition.2008.07.014>
- Krajbich, I., Armel, C., & Rangel, A. (2010). Visual fixations and the computation and comparison of value in simple choice. *Nature Neuroscience*, *13*(10), 1292–1298. <https://doi.org/10.1038/nn.2635>
- Kunert, R. (2016). Internal conceptual replications do not increase independent replication success. *Psychonomic Bulletin & Review*. <https://doi.org/10.3758/s13423-016-1030-9>
- Kurzban, R., Duckworth, A., Kable, J. W., & Myers, J. (2013). An opportunity cost model of subjective effort and task performance. *Behavioral and Brain Sciences*, *36*(06), 661–679. <https://doi.org/10.1017/S0140525X12003196>
- Kushlev, K., & Dunn, E. W. (2015). Checking email less frequently reduces stress. *Computers in Human Behavior*, *43*, 220–228. <https://doi.org/10.1016/j.chb.2014.11.005>
- Kushlev, K., Hunter, J. F., Proulx, J., Pressman, S. D., & Dunn, E. (2019). Smartphones reduce smiles between strangers. *Computers in Human Behavior*, *91*, 12–16. <https://doi.org/10.1016/j.chb.2018.09.023>
- Kushlev, K., Proulx, J., & Dunn, E. W. (2016). “Silence your phones”: Smartphone notifications increase inattention and hyperactivity symptoms. *Proceedings of the 2016 CHI*

- Conference on Human Factors in Computing Systems - CHI '16*, 1011–1020. <https://doi.org/10.1145/2858036.2858359>
- Lakens, D., & Evers, E. R. K. (2014). Sailing from the seas of chaos into the corridor of stability: Practical recommendations to increase the informational value of studies. *Perspectives on Psychological Science*, 9(3), 278–292. <https://doi.org/10.1177/1745691614528520>
- Lakens, D., Scheel, A. M., & Isager, P. M. (2018). Equivalence testing for psychological research: A tutorial. *Advances in Methods and Practices in Psychological Science*, 1(2), 259–269.
- LaRose, R. (2015). The psychology of interactive media habits. In S. S. Sundar (Ed.), *The Handbook of the Psychology of Communication Technology* (pp. 365–383). Chichester, UK: John Wiley & Sons, Inc.
- Lawrence, N. S., O'Sullivan, J., Parslow, D., Javaid, M., Adams, R. C., Chambers, C. D., ... Verbruggen, F. (2015). Training response inhibition to food is associated with weight loss and reduced energy intake. *Appetite*, 95, 17–28. <https://doi.org/10.1016/j.appet.2015.06.009>
- Le Pelley, M. E., Mitchell, C. J., Beesley, T., George, D. N., & Wills, A. J. (2016). Attention and associative learning in humans: An integrative review. *Psychological Bulletin*, 142(10), 1111–1140. <https://doi.org/10.1037/bul0000064>
- Leary, M. R., Kelly, K. M., Cottrell, C. A., & Schreindorfer, L. S. (2007). Individual differences in the need to belong: Mapping the nomological network. *Unpublished Manuscript, Duke University*.
- Lee, M. D., & Wagenmakers, E.-J. (2013). *Bayesian cognitive modeling: A practical course*. New York, NY: Cambridge University Press.
- Leiner, D. J. (2013). Too Fast, too straight, too weird: Post hoc identification of meaningless data in internet surveys. <https://doi.org/10.2139/ssrn.2361661>
- Levine, L. E., Waite, B. M., & Bowman, L. L. (2013). Use of instant messaging predicts self-report but not performance measures of inattention, impulsiveness, and distractibility. *Cyberpsychology, Behavior and Social Networking*, 16(12), 898–903. <https://doi.org/10.1089/cyber.2012.0504>
- Logan, G. D. (1994). On the ability to inhibit thought and action: A user's guide to the stop signal paradigm. In D. Dagenbach & T. H. Carr (Eds.), *Inhibitory processes in attention, memory, and language* (pp. 189–239). San Diego, CA: Academic Press.
- Luke, S. G. (2017). Evaluating significance in linear mixed-effects models in R. *Behavior Research Methods*, 49(4), 1494–1502. <https://doi.org/10.3758/s13428-016-0809-y>
- Marino, C., Gini, G., Vieno, A., & Spada, M. M. (2018). A comprehensive meta-analysis on problematic facebook use. *Computers in Human Behavior*, 83, 262–277. <https://doi.org/10.1016/j.chb.2018.02.009>
- Marteau, T. M., Hollands, G. J., & Fletcher, P. C. (2012). Changing human behavior to prevent disease: The importance of targeting automatic processes. *Science*, 337(6101), 1492–1495. <https://doi.org/10.1126/science.1226918>
- Mascheroni, G., & Vincent, J. (2016). Perpetual contact as a communicative affordance: Opportunities, constraints, and emotions. *Mobile Media & Communication*, 4(3), 310–326. <https://doi.org/10.1177/2050157916639347>

- Masur, P. K. (2018). *Situational privacy and self-disclosure: Communication processes in online environments*. Cham, Switzerland: Springer.
- McEwan, B., Carpenter, C. J., & Westerman, D. (2018). On replication in Communication Science. *Communication Studies*, 69(3), 235–241. <https://doi.org/10.1080/10510974.2018.1464938>
- Mehrotra, A., Pejovic, V., Vermeulen, J., Hendley, R., & Musolesi, M. (2016). My phone and me: Understanding people's receptivity to mobile notifications. *Proceedings of the 2016 CHI Conference on Human Factors in Computing Systems*, 1021–1032. <https://doi.org/10.1145/2858036.2858566>
- Meier, A., Domahidi, E., & Günter, E. (in press). Computer-mediated communication and mental health: A computational scoping review of an interdisciplinary field. In S. Yates & R. E. Rice (Eds.), *The Oxford handbook of digital technology and society*. Oxford, United Kingdom: Oxford University Press.
- Meier, A., Reinecke, L., & Meltzer, C. E. (2016). "Facebocrastination"? Predictors of using Facebook for procrastination and its effects on students' well-being. *Computers in Human Behavior*, 64, 65–76. <https://doi.org/10.1016/j.chb.2016.06.011>
- Mercer-Lynn, K. B., Flora, D. B., Fahlman, S. A., & Eastwood, J. D. (2013). The measurement of boredom: Differences between existing self-report scales. *Assessment*, 20(5), 585–596. <https://doi.org/10.1177/1073191111408229>
- Mihailidis, P. (2014). A tethered generation: Exploring the role of mobile phones in the daily life of young people. *Mobile Media & Communication*, 2(1), 58–72. <https://doi.org/10.1177/2050157913505558>
- Miller, G. (2012). The Smartphone Psychology Manifesto. *Perspectives on Psychological Science*, 7(3), 221–237. <https://doi.org/10.1177/1745691612441215>
- Miyake, A., & Friedman, N. P. (2012). The nature and organization of individual differences in executive functions: Four general conclusions. *Current Directions in Psychological Science*, 21(1), 8–14. <https://doi.org/10.1177/0963721411429458>
- Mohan, M., Khaliq, F., Panwar, A., & Vaney, N. (2016). Does chronic exposure to mobile phones affect cognition? *Functional Neurology*, 31(1), 47–51.
- Mooneyham, B. W., & Schooler, J. W. (2013). The costs and benefits of mind-wandering: A review. *Canadian Journal of Experimental Psychology/Revue Canadienne de Psychologie Expérimentale*, 67(1), 11–18. <https://doi.org/10.1037/a0031569>
- Morey, R. D. (2008). Confidence intervals from normalized data: A correction to Cousineau (2005). *Tutorials in Quantitative Methods for Psychology*, 4(2), 61–64. <https://doi.org/10.20982/tqmp.04.2.p061>
- Morey, R. D., & Rouder, J. N. (2015). *BayesFactor: Computation of Bayes Factors for common designs. R package version 0.9.12-2*. Retrieved from <https://cran.r-project.org/package=BayesFactor>
- Mrazek, M. D., Smallwood, J., & Schooler, J. W. (2012). Mindfulness and mind-wandering: Finding convergence through opposing constructs. *Emotion*, 12(3), 442–448. <https://doi.org/10.1037/a0026678>

- Munafò, M. R., Nosek, B. A., Bishop, D. V. M., Button, K. S., Chambers, C. D., Percie, N., ... Wagenmakers, E.-J. (2017). A manifesto for reproducible science. *Human Nature Behavior*, *1*, 1–9. <https://doi.org/10.1038/s41562-016-0021>
- Näsi, M., & Koivusilta, L. (2013). Internet and everyday life: The perceived implications of internet use on memory and ability to concentrate. *CyberPsychology, Behavior, & Social Networking*, *16*(2), 88–93. <https://doi.org/10.1089/cyber.2012.0058>
- Nelson, L. D., & Simmons, J. (2018). Psychology's Renaissance. *Annual Review of Psychology*, *69*, 511–534.
- Nosek, B. A., Ebersole, C. R., DeHaven, A. C., & Mellor, D. T. (2018). The preregistration revolution. *Proceedings of the National Academy of Sciences*, *115*(11), 2600–2606. <https://doi.org/10.1073/pnas.1708274114>
- Nosek, B. A., & Lakens, D. (2014). Registered reports: A method to increase the credibility of published results. *Social Psychology*, *45*(3), 137–141. <https://doi.org/10.1027/1864-9335/a000192>
- Okeke, F., Sobolev, M., Dell, N., & Estrin, D. (2018). Good vibrations: Can a digital nudge reduce digital overload? *Proceedings of the 20th International Conference on Human-Computer Interaction with Mobile Devices and Services*, 4:1–4:12. <https://doi.org/10.1145/3229434.3229463>
- Open Science Collaboration. (2015). Estimating the reproducibility of psychological science. *Science*, *349*(6251), aac4716–aac4716. <https://doi.org/10.1126/science.aac4716>
- Orben, A., Dienlin, T., & Przybylski, A. K. (2019). Social media's enduring effect on adolescent life satisfaction. *Proceedings of the National Academy of Sciences*, 201902058. <https://doi.org/10.1073/pnas.1902058116>
- Orben, A., & Przybylski, A. K. (2019a). Screens, teens, and psychological well-being: Evidence from three time-use-diary studies. *Psychological Science*, 095679761983032. <https://doi.org/10.1177/0956797619830329>
- Orben, A., & Przybylski, A. K. (2019b). The association between adolescent well-being and digital technology use. *Nature Human Behaviour*. <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>
- Panek, E. (2014). Left to their own devices: College students' "guilty pleasure" media use and time management. *Communication Research*, *41*(4), 561–577. <https://doi.org/10.1177/0093650213499657>
- Papies, E. K. (2013). Tempting food words activate eating simulations. *Frontiers in Psychology*, *4*, 1–12. <https://doi.org/10.3389/fpsyg.2013.00838>
- Papies, E. K., & Barsalou, L. W. (2015). Grounding desire and motivated behavior: A theoretical framework and review of empirical evidence. In W. Hofmann & L. F. Nordgren (Eds.), *The psychology of desire* (p. 21). New York: Guilford Press.
- Parry, D. A., & le Roux, D. B. (2019). Media multitasking and cognitive control: A systematic review of interventions. *Computers in Human Behavior*, *92*, 316–327. <https://doi.org/10.1016/j.chb.2018.11.031>

- Peirce, J. W. (2007). PsychoPy—Psychophysics software in Python. *Journal of Neuroscience Methods*, 162(1–2), 8–13.
- Pew Research Center. (2015). U.S. smartphone use in 2015. Retrieved from <http://www.pewinternet.org/2015/04/01/us-smartphone-use-in-2015/>
- Pew Research Center. (2017). Mobile fact sheet. Retrieved from <http://www.pewinternet.org/fact-sheet/mobile/>
- Pew Research Center. (2018, August 22). How teens and parents navigate screen time and device distractions. Retrieved June 17, 2019, from <https://www.pewinternet.org/2018/08/22/how-teens-and-parents-navigate-screen-time-and-device-distractions/>
- Pielot, M., Church, K., & de Oliveira, R. (2014). An in-situ study of mobile phone notifications. *Proceedings of the 16th International Conference on Human-Computer Interaction with Mobile Devices & Services - MobileHCI '14*, 233–242. <https://doi.org/10.1145/2628363.2628364>
- Pielot, M., de Oliveira, R., Kwak, H., & Oliver, N. (2014). Didn't you see my message? Predicting attentiveness to mobile instant messages. *Proceedings of the 32nd Annual ACM Conference on Human Factors in Computing Systems - CHI '14*, 3319–3328. <https://doi.org/10.1145/2556288.2556973>
- Poerio, G. L., Totterdell, P., Emerson, L.-M., & Miles, E. (2016). Social daydreaming and adjustment: An experience-sampling study of socio-emotional adaptation during a life transition. *Frontiers in Psychology*, 7. <https://doi.org/10.3389/fpsyg.2016.00013>
- Pool, E., Brosch, T., Delplanque, S., & Sander, D. (2016). Attentional bias for positive emotional stimuli: A meta-analytic investigation. *Psychological Bulletin*, 142(1), 79.
- Postman, N. (1985). *Amusing ourselves to death: Public discourse in the age of show business*. Viking Penguin.
- Potter, W. J. (2011). Conceptualizing mass media effect. *Journal of Communication*, 61(5), 896–915. <https://doi.org/10.1111/j.1460-2466.2011.01586.x>
- Przybylski, A. K., Murayama, K., DeHaan, C. R., & Gladwell, V. (2013). Motivational, emotional, and behavioral correlates of fear of missing out. *Computers in Human Behavior*, 29, 1841–1848. <https://doi.org/10.1016/j.chb.2013.02.014>
- Przybylski, A. K., & Weinstein, N. (2013). Can you connect with me now? How the presence of mobile communication technology influences face-to-face conversation quality. *Journal of Social and Personal Relationships*, 30(3), 237–246. <https://doi.org/10.1177/0265407512453827>
- Przybylski, A. K., & Weinstein, N. (2017). A large-scale test of the Goldilocks Hypothesis: Quantifying the relations between digital-screen use and the mental well-being of adolescents. *Psychological Science*, 1–12. <https://doi.org/10.1177/0956797616678438>
- Python Core Team. (2018). *Python: A dynamic, open source programming language*. Python Software Foundation. Retrieved from <https://www.python.org/>
- R Core Team. (2018). *R: A language and environment for statistical computing*. R Foundation for Statistical Computing, Vienna, Austria. Retrieved from <https://www.r-project.org/>
- Rains, S. A., Levine, T. R., & Weber, R. (2018). Sixty years of quantitative communication research summarized: Lessons from 149 meta-analyses. *Annals of the International Communication Association*, 8985, 1–20. <https://doi.org/10.1080/23808985.2018.1446350>

- Reich, S., Schneider, F. M., & Heling, L. (2018). Zero Likes – Symbolic interactions and need satisfaction online. *Computers in Human Behavior, 80*, 97–102. <https://doi.org/10.1016/j.chb.2017.10.043>
- Reinecke, L. (2018). POPC and well-being: A risk-benefit analysis. In P. Vorderer, D. Hefner, L. Reinecke, & C. Klimmt (Eds.), *Permanently online, permanently connected. Living and communication in a POPC world* (pp. 233–243). New York: Routledge.
- Reinecke, L., Aufenanger, S., Beutel, M. E., Dreier, M., Quiring, O., Stark, B., ... Müller, K. W. (2017). Digital stress over the life span: The effects of communication load and internet multitasking on perceived stress and psychological health impairments in a German probability sample. *Media Psychology, 20*, 90–115. <https://doi.org/10.1080/15213269.2015.1121832>
- Reinecke, L., Hartmann, T., & Eden, A. (2014). The guilty couch potato: The role of ego depletion in reducing recovery through media use. *Journal of Communication, 64*, 569–589. <https://doi.org/10.1111/jcom.12107>
- 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., ... 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>
- Reinecke, L., Klimmt, C., Meier, A., Reich, S., Hefner, D., Knop-Hülß, K., ... Vorderer, P. (2017). Development of the online vigilance scale. *67th Annual Conference of the International Communication Association*. Presented at the San Diego, CA. San Diego, CA.
- Roberts, J. A., & Pirog, S. F. (2013). A preliminary investigation of materialism and impulsiveness as predictors of technological addictions among young adults. *Journal of Behavioral Addictions, 2*(1), 56–62. <https://doi.org/10.1556/JBA.1.2012.011>
- Rosseel, Y. (2012). lavaan: An R package for structural equation modeling. *Journal of Statistical Software, 48*(2), 1–36.
- Rusz, D., Bijleveld, E., & Kompier, M. A. J. (2018). Reward-associated distractors can harm cognitive performance. *PLOS ONE, 13*(10), e0205091. <https://doi.org/10.1371/journal.pone.0205091>
- Salomon, T., Botvinik-Nezer, R., Gutentag, T., Gera, R., Iwanir, R., Tamir, M., & Schonberg, T. (2018). The cue-approach task as a general mechanism for long-term non-reinforced behavioral change. *Scientific Reports, 8*(1), 3614. <https://doi.org/10.1038/s41598-018-21774-3>
- Sanbonmatsu, D. M., Strayer, D. L., Medeiros-Ward, N., & Watson, J. M. (2013). Who multi-tasks and why? Multi-tasking ability, perceived multi-tasking ability, impulsivity, and sensation seeking. *PLoS ONE, 8*(1), e54402. <https://doi.org/10.1371/journal.pone.0054402>
- Santor, D. A., Messervey, D., & Kusumakar, V. (2000). Measuring peer pressure, popularity, and conformity in adolescent boys and girls: Predicting school performance, sexual attitudes, and substances. *Journal of Youth and Adolescence, 29*(2), 163–182.
- Sauer, S., Walach, H., Schmidt, S., Hinterberger, T., Lynch, S., Büssing, A., & Kohls, N. (2013). Assessment of mindfulness: Review on state of the art. *Mindfulness, 4*(1), 3–17. <https://doi.org/10.1007/s12671-012-0122-5>

- Saunders, B., Milyavskaya, M., Etz, A., Randles, D., & Inzlicht, M. (2018). Reported self-control is not meaningfully associated with inhibition-related executive function: A bayesian analysis. *Collabra: Psychology*, 4(1), 39. <https://doi.org/10.1525/collabra.134>
- Schäfer, T., & Schwarz, M. A. (2019). The meaningfulness of effect sizes in psychological research: Differences between sub-disciplines and the impact of potential biases. *Frontiers in Psychology*, 10. <https://doi.org/10.3389/fpsyg.2019.00813>
- 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>
- Scherbaum, C. a., & Ferrerter, J. M. (2008). Estimating statistical power and required sample sizes for organizational research using multilevel modeling. *Organizational Research Methods*, 12(2), 347–367. <https://doi.org/10.1177/1094428107308906>
- Scholten, H., Granic, I., Chen, Z., Veling, H., & Luijten, M. (2019). Do smokers devalue smoking cues after go/no-go training? *Psychology & Health*, 1–17. <https://doi.org/10.1080/08870446.2018.1554184>
- Schonberg, T., Bakkour, A., Hover, A. M., Mumford, J. A., Nagar, L., Perez, J., & Poldrack, R. A. (2014). Changing value through cued approach: An automatic mechanism of behavior change. *Nature Neuroscience*, 17(4), 625–630. <https://doi.org/10.1038/nn.3673>
- Schönbrodt, F. D., Wagenmakers, E.-J., Zehetleitner, M., & Perugini, M. (2017). Sequential hypothesis testing with Bayes factors: Efficiently testing mean differences. *Psychological Methods*, 22(2), 322–339. <https://doi.org/10.1037/met0000061>
- Seibt, B., Häfner, M., & Deutsch, R. (2007). Prepared to eat: How immediate affective and motivational responses to food cues are influenced by food deprivation. *European Journal of Social Psychology*, 37(2), 359–379. <https://doi.org/10.1002/ejsp.365>
- Seli, P., Kane, M. J., Smallwood, J., Schacter, D. L., Maillet, D., Schooler, J. W., & Smilek, D. (2018). Mind-Wandering as a natural kind: A family-resemblances view. *Trends in Cognitive Sciences*, 22(6), 479–490. <https://doi.org/10.1016/j.tics.2018.03.010>
- Seo, M., Kim, J.-H., & David, P. (2015). Always connected or always distracted? ADHD symptoms and social assurance explain problematic use of mobile phone and multicomputing. *Journal of Computer-Mediated Communication*, 20, 667–681. <https://doi.org/10.1111/jcc4.12140>
- Sheeran, P., Maki, A., Montanaro, E., Avishai-Yitshak, A., Bryan, A., Klein, W. M. P., ... Rothman, A. J. (2016). The impact of changing attitudes, norms, and self-efficacy on health-related intentions and behavior: A meta-analysis. *Health Psychology*, 35(11), 1178–1188. <https://doi.org/10.1037/hea0000387>
- Sheeran, P., & Webb, T. L. (2016). The intention–behavior gap. *Social and Personality Psychology Compass*, 10(9), 503–518.
- Shelton, J. T., Elliott, E. M., Eaves, S. D., & Exner, A. L. (2009). The distracting effects of a ringing cell phone: An investigation of the laboratory and the classroom setting. *Journal of Environmental Psychology*, 29(4), 513–521. <https://doi.org/10.1016/j.jenvp.2009.03.001>

- Shin, J., & Shin, M. (2016). To be connected or not to be connected? Mobile messenger overload, fatigue, and mobile shunning. *Cyberpsychology, Behavior, and Social Networking, 19*(10), 579–586. <https://doi.org/10.1089/cyber.2016.0236>
- Shrout, P. E., & Bolger, N. (2002). Mediation in experimental and nonexperimental studies: New procedures and recommendations. *Psychological Methods, 7*(4), 422–445. <https://doi.org/10.1037/1082-989X.7.4.422>
- Simmons, J. P., Nelson, L. D., & Simonsohn, U. (2011). False-positive psychology: Undisclosed flexibility in data collection and analysis allows presenting anything as significant. *Psychological Science, 22*(11), 1359–1366. <https://doi.org/10.1177/0956797611417632>
- Singmann, H., Bolker, B., Westfall, J., & Aust, F. (2018). *Afex: Analysis of factorial experiments. R package version 0.20-2*. Retrieved from <https://cran.r-project.org/package=afex>
- Smallwood, J., O'Connor, R. C., Sudbery, M. V., & Obonsawin, M. (2007). Mind-wandering and dysphoria. *Cognition and Emotion, 21*(4), 816–842. <https://doi.org/10.1080/02699930600911531>
- Smallwood, J., & Schooler, J. W. (2015). The science of mind wandering: Empirically navigating the stream of consciousness. *Annual Review of Psychology, 66*(1), 487–518. <https://doi.org/10.1146/annurev-psych-010814-015331>
- Smetaniuk, P. (2014). A preliminary investigation into the prevalence and prediction of problematic cell phone use. *Journal of Behavioral Addictions, 3*(1), 41–53. <https://doi.org/10.1556/JBA.3.2014.004>
- Smith, T. S., Isaak, M. I., Senette, C. G., & Abadie, B. G. (2011). Effects of cell-phone and text-message distractions on true and false recognition. *Cyberpsychology Behavior and Social Networking, 14*(6), 351–358. <https://doi.org/10.1089/cyber.2010.0129>
- Soror, A. A., Hammer, B. I., Steelman, Z. R., Davis, F. D., & Limayem, M. M. (2015). Good habits gone bad: Explaining negative consequences associated with the use of mobile phones from a dual-systems perspective. *Information Systems Journal, 25*, 403–427. <https://doi.org/10.1111/isj.12065>
- Spronken, M., Holland, R. W., Figner, B., & Dijksterhuis, A. (2016). Temporal focus, temporal distance, and mind-wandering valence: Results from an experience sampling and an experimental study. *Consciousness and Cognition, 41*, 104–118. <https://doi.org/10.1016/j.concog.2016.02.004>
- Sriwilai, K., & Charoensukmongkol, P. (2016). Face it, don't Facebook it: Impacts of social media addiction on mindfulness, coping strategies and the consequence on emotional exhaustion. *Stress and Health, 32*(4), 427–434. <https://doi.org/10.1002/smi.2637>
- Stawarczyk, D., Majerus, S., der Linden, M. V., & D'Argembeau, A. (2012). Using the daydreaming frequency scale to investigate the relationships between mind-wandering, psychological well-being, and present-moment awareness. *Frontiers in Psychology, 3*, 1–15. <https://doi.org/10.3389/fpsyg.2012.00363>
- Stice, E., Lawrence, N. S., Kemps, E., & Veling, H. (2016). Training motor responses to food: A novel treatment for obesity targeting implicit processes. *Clinical Psychology Review, 49*, 16–27. <https://doi.org/10.1016/j.cpr.2016.06.005>

- Stothart, C., Mitchum, A., & Yehnert, C. (2015). The attentional cost of receiving a cell phone notification. *Journal of Experimental Psychology: Human Perception and Performance*, 41(4), 893–897. <https://doi.org/10.1037/xhp0000100>
- Strobach, T., Liepelt, R., Schubert, T., & Kiesel, A. (2012). Task switching: Effects of practice on switch and mixing costs. *Psychological Research*, 76(1), 74–83. <https://doi.org/10.1007/s00426-011-0323-x>
- Süddeutsche Zeitung. (2015). "Smombie" ist das Jugendwort des Jahres ["Smombie" is youth word of the year]. Retrieved from <http://www.sueddeutsche.de/kultur/deutsche-sprache-smombie-ist-das-jugendwort-des-jahres-1.2735599>
- Sundar, S. S., & Limperos, A. M. (2013). Uses and grats 2.0: New gratifications for new media. *Journal of Broadcasting & Electronic Media*, 57(4), 504–525. <https://doi.org/10.1080/08838151.2013.845827>
- Tanis, M., Beukeboom, C. J., Hartmann, T., & Vermeulen, I. E. (2015). Phantom phone signals: An investigation into the prevalence and predictors of imagined cell phone signals. *Computers in Human Behavior*, 51, 356–362. <https://doi.org/10.1016/j.chb.2015.04.039>
- Thaler, R. H., & Sunstein, C. R. (2009). *Nudge: Improving decisions about health, wealth, and happiness*. Penguin.
- Theeuwes, J., & Belopolsky, A. V. (2012). Reward grabs the eye: Oculomotor capture by rewarding stimuli. *Vision Research*, 74, 80–85. <https://doi.org/10.1016/j.visres.2012.07.024>
- Thomson, D. R., Besner, D., & Smilek, D. (2015). A resource-control account of sustained attention: Evidence from mind wandering and vigilance paradigms. *Perspectives on Psychological Science*, 10(1), 82–96. <https://doi.org/10.1177/1745691614556681>
- Thornton, B., Faires, A., Robbins, M., & Rollins, E. (2014). The mere presence of a cell phone may be distracting. *Social Psychology*, 45(6), 479–488. <https://doi.org/10.1027/1864-9335/a000216>
- Trepte, S., Dienlin, T., & Reinecke, L. (2015). Influence of social support received in online and offline contexts on satisfaction with social support and satisfaction with life: A longitudinal study. *Media Psychology*, 18(1), 74–105. <https://doi.org/10.1080/15213269.2013.838904>
- Tuk, M. A., Zhang, K., & Sweldens, S. (2015). The propagation of self-control: Self-control in one domain simultaneously improves self-control in other domains. *Journal of Experimental Psychology: General*, 144(3), 639–654. <https://doi.org/10.1037/xge0000065>
- Turkle, S. (2012). *Alone together: Why we expect more from technology and less from each other*. Basic books.
- Turton, R., Bruidegom, K., Cardi, V., Hirsch, C. R., & Treasure, J. (2016). Novel methods to help develop healthier eating habits for eating and weight disorders: A systematic review and meta-analysis. *Neuroscience & Biobehavioral Reviews*, 61, 132–155. <https://doi.org/10.1016/j.neubiorev.2015.12.008>
- Valkenburg, P. M., & Peter, J. (2013). The differential susceptibility to media effects model. *Journal of Communication*, 63(2), 221–243. <https://doi.org/10.1111/jcom.12024>

- van der Schuur, W. A., Baumgartner, S. E., Sumter, S. R., & Valkenburg, P. M. (2015). The consequences of media multitasking for youth: A review. *Computers in Human Behavior, 53*, 204–215. <https://doi.org/10.1016/j.chb.2015.06.035>
- Van Dessel, P., Eder, A. B., & Hughes, S. (2018). Mechanisms underlying effects of approach-avoidance training on stimulus evaluation. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 44*(8), 1224–1241. <https://doi.org/10.1037/xlm0000514>
- Van Dessel, P., Hughes, S., & De Houwer, J. (2018a). Consequence-based approach-avoidance training: A new and improved method for changing behavior. *Psychological Science, 29*(12), 1899–1910. <https://doi.org/10.1177/0956797618796478>
- Van Dessel, P., Hughes, S., & De Houwer, J. (2018b). How do actions influence attitudes? An inferential account of the impact of action performance on stimulus evaluation. *Personality and Social Psychology Review, 1*–18. <https://doi.org/10.1177/1088868318795730>
- van Koningsbruggen, G. M., Hartmann, T., Eden, A., & Veling, H. (2017). Spontaneous hedonic reactions to social media cues. *Cyberpsychology, Behavior, and Social Networking, 20*(5), 334–340. <https://doi.org/10.1089/cyber.2016.0530>
- 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>
- Vasishth, S., Mertzen, D., Jäger, L. A., & Gelman, A. (2018). The statistical significance filter leads to overoptimistic expectations of replicability. *Journal of Memory and Language, 103*, 151–175. <https://doi.org/10.1016/j.jml.2018.07.004>
- Veling, H., Aarts, H., & Stroebe, W. (2013). Stop signals decrease choices for palatable foods through decreased food evaluation. *Frontiers in Psychology, 4*. <https://doi.org/10.3389/fpsyg.2013.00875>
- Veling, H., Lawrence, N. S., Chen, Z., van Koningsbruggen, G. M., & Holland, R. W. (2017). What is trained during food go/no-go training? A review focusing on mechanisms and a research agenda. *Current Addiction Reports, 4*(1), 35–41. <https://doi.org/10.1007/s40429-017-0131-5>
- Verbruggen, F., Aron, A. R., Band, G., Beste, C., Bissett, P., Brockett, A. T., ... Boehler, C. N. (2019). *Capturing the ability to inhibit actions and impulsive behaviors: A consensus guide to the stop-signal task*. <https://doi.org/10.31219/osf.io/8mzdu>
- Verbruggen, F., Aron, A. R., Stevens, M. A., & Chambers, C. D. (2010). Theta burst stimulation dissociates attention and action updating in human inferior frontal cortex. *Proceedings of the National Academy of Sciences of the United States of America, 107*(31), 13966–13971. <https://doi.org/10.1073/pnas.1001957107>
- Verbruggen, F., Chambers, C. D., & Logan, G. D. (2013). Fictitious inhibitory differences: How skewness and slowing distort the estimation of stopping latencies. *Psychological Science, 24*(3), 352–362. <https://doi.org/10.1177/0956797612457390>
- Verbruggen, F., & Logan, G. D. (2008a). Automatic and controlled response inhibition: Associative learning in the go/no-go and stop-signal paradigms. *Journal of Experimental Psychology: General, 137*(4), 649–672. <https://doi.org/10.1037/a0013170>
- Verbruggen, F., & Logan, G. D. (2008b). Response inhibition in the stop-signal paradigm. *Trends in Cognitive Sciences, 12*(11), 418–424. <https://doi.org/10.1016/j.tics.2008.07.005>

- Verbruggen, F., & Logan, G. D. (2009). Proactive adjustments of response strategies in the stop-signal paradigm. *Journal of Experimental Psychology: Human Perception and Performance*, 35(3), 835–854. <https://doi.org/10.1037/a0012726>
- Verbruggen, F., McLaren, I. P. L., & Chambers, C. D. (2014). Banishing the control homunculi in studies of action control and behavior change. *Perspectives on Psychological Science*, 9(5), 497–524. <https://doi.org/10.1177/1745691614526414>
- Verbruggen, F., Stevens, T., & Chambers, C. D. (2014). Proactive and reactive stopping when distracted: An attentional account. *Journal of Experimental Psychology: Human Perception and Performance*, 40(4), 1295–1300. <https://doi.org/10.1037/a0036542>
- Verduyn, P., Lee, D. S., Park, J., Shablack, H., Orvell, A., Bayer, J., ... Kross, E. (2015). Passive Facebook usage undermines affective well-being: Experimental and longitudinal evidence. *Journal of Experimental Psychology: General*, 144(2), 480–488. <https://doi.org/10.1037/xge0000057>
- Vermeulen, I., & Hartmann, T. (2015). Questionable research and publication practices in communication science. *Communication Methods and Measures*, 9(4), 189–192. <https://doi.org/10.1080/19312458.2015.1096331>
- Verplanken, B., & Orbell, S. (2003). Reflections on past behavior: A self-report index of habit strength. *Journal of Applied Social Psychology*, 33(6), 1313–1330. <https://doi.org/10.1111/j.1559-1816.2003.tb01951.x>
- Vorderer, P., & Kohring, M. (2013). Permanently online: A challenge for media and communication research. *International Journal of Communication*, 7, 188–196. <https://doi.org/1932-8036/2013FEA0002>
- Wagenmakers, E.-J. (2007). A practical solution to the pervasive problems of p values. *Psychonomic Bulletin & Review*, 14(5), 779–804. <https://doi.org/10.3758/BF03194105>
- Wagenmakers, E.-J., Marsman, M., Jamil, T., Ly, A., Verhagen, J., Love, J., ... Morey, R. D. (2018). Bayesian inference for psychology. Part I: Theoretical advantages and practical ramifications. *Psychonomic Bulletin and Review*, 25(1), 35–57. <https://doi.org/10.3758/s13423-017-1343-3>
- Wagenmakers, E.-J., Wetzels, R., Borsboom, D., van der Maas, H. L. J., & Kievit, R. A. (2012). An agenda for purely confirmatory research. *Perspectives on Psychological Science*, 7, 627–633. <https://doi.org/10.1177/1745691612463078>
- Walther, J. B. (2017). The merger of mass and interpersonal communication via new media: Integrating metaconstructs. *Human Communication Research*, 1–14. <https://doi.org/10.1111/hcre.12122>
- Ward, A. F., Duke, K., Gneezy, A., & Bos, M. W. (2017). Brain drain: The mere presence of one's own smartphone reduces available cognitive capacity. *Journal of the Association for Consumer Research*, 2(2), 140–154. <https://doi.org/10.1086/691462>
- Warm, J. S., Finomore, V. S., Vidulich, M. A., & Funke, M. E. (2015). Vigilance: A perceptual challenge. In *Cambridge Handbooks in Psychology. The Cambridge handbook of applied perception research, Vol. I.* (pp. 241–283). <https://doi.org/10.1017/CBO9780511973017.018>
- Warm, J. S., Matthews, G., & Finomore Jr, V. S. (2018). Vigilance, workload, and stress. In *Performance under stress* (pp. 131–158). CRC Press.

- 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>.
- Wickham, H. (2017). tidyverse: Easily install and load the “tidyverse” (Version 1.2.1.). Retrieved from <https://CRAN.R-project.org/package=tidyverse>
- Wiers, R. W., Eberl, C., Rinck, M., Becker, E. S., & Lindenmeyer, J. (2011). Retraining automatic action tendencies changes alcoholic patients’ approach bias for alcohol and improves treatment outcome. *Psychological Science*, *22*(4), 490–497. <https://doi.org/10.1177/0956797611400615>
- Wilcockson, T. D. W., Ellis, D. A., & Shaw, H. (2018). Determining typical smartphone usage: What data do we need? *Cyberpsychology, Behavior, and Social Networking*, *00*(00), cyber.2017.0652. <https://doi.org/10.1089/cyber.2017.0652>
- Wilhelm, P., & Schoebi, D. (2007). Assessing mood in daily life: Structural validity, sensitivity to change, and reliability of a short-scale to measure three basic dimensions of mood. *European Journal of Psychological Assessment*, *23*(4), 258–267. <https://doi.org/10.1027/1015-5759.23.4.258>
- Wilmer, H. H., & Chein, J. M. (2016). Mobile technology habits: Patterns of association among device usage, intertemporal preference, impulse control, and reward sensitivity. *Psychonomic Bulletin & Review*, 1–8. <https://doi.org/10.3758/s13423-016-1011-z>
- Zoltak, M. J., Veling, H., Chen, Z., & Holland, R. W. (2018). Attention! Can choices for low value food over high value food be trained? *Appetite*, *124*, 124–132. <https://doi.org/10.1016/j.appet.2017.06.010>

Samenvatting

Smartphones zijn overal en veel mensen maken zich zorgen dat ze een nadelig effect hebben op onze prestaties en ons welzijn. Er is aangetoond dat *smartphonesignalen*, zoals meldingen of zelfs de aanwezigheid van een telefoon, afleidend kunnen zijn. Bovendien kunnen deze signalen leiden tot een voortdurend alerte gemoedstoestand, ook wel *online waakzaamheid* genoemd. Gebruikers ervaren zowel smartphonesignalen als de daaruit voortvloeiende online waakzaamheid als hinderlijk. Er is echter weinig onderzoek gedaan naar de vraag of de bezorgdheid over negatieve effecten op prestaties en welzijn al dan niet gerechtvaardigd is. Daarom heeft dit proefschrift het algemene doel om de effecten van smartphonesignalen en online waakzaamheid op welzijn en prestaties te onderzoeken.

Voor de beantwoording van deze vraag stelden we onszelf vier doelen. Allereerst onderzochten we het verband tussen online waakzaamheid en welzijn op het niveau van personeigenschappen (hoofdstuk 2) en op het niveau van dagelijkse omstandigheden (hoofdstuk 3). Ten tweede toetsten we of smartphonesignalen een verstoring van cognitieve inhibitie (hoofdstuk 4) en het werkgeheugen (hoofdstuk 5) tot gevolg hadden. Ten derde hebben we onderzocht of online waakzaamheid kan worden verklaard aan de hand van de belonende aard van smartphones (hoofdstuk 6). Ten vierde hebben we getest of we de voorkeur van personen voor smartphonesignalen kunnen reduceren (hoofdstuk 7).

DOEL 1: IS ER EEN VERBAND TUSSEN ONLINE WAAKZAAMHEID EN WELZIJN?

In hoofdstuk 2 hebben we getest of er een negatief verband is tussen online waakzaamheid en welzijn, wat wordt geuit in de vorm van absentia mentalis, en in het bijzonder van dwalende gedachten (*mindwandering*) en verminderde mindfulness. Om een vollediger beeld te krijgen van mogelijke verbanden met welzijn, hebben we geëvalueerd hoe proefpersonen hun leven beoordelen (oftewel de cognitieve component van welzijn) en hoe zij zich meestal voelen (oftewel de affectieve component van welzijn). Zoals voorspeld, waren online waakzaamheid en de twee welzijnscomponenten indirect aan elkaar verbonden: Online waakzaamheid bleek aan de welzijnsresultaten verbonden te zijn door een verminderd niveau van mindfulness, maar niet door dwalende gedachten.

Hoofdstuk 3 is gebaseerd op een combinatie van in situ zelfrapportages en objectief smartphonegebruik om de relatie tussen online waakzaamheid en welzijn in het dagelijks leven te beoordelen. Onze resultaten tonen aan dat online waakzaamheid, indien geuit door gedragingen als het controleren van de telefoon of het reageren op meldingen, niet aan welzijn is verbonden. Wanneer online waakzaamheid werd geuit in de vorm van gedachten over onlinecommunicatie, was online waakzaamheid op negatieve wijze aan welzijn gerelateerd. Hoe vaker de deelnemers in het afgelopen half uur aan online interacties dachten, hoe slechter ze zich op dat moment voelden. Dit verband was echter klein en de vraag of dergelijke gedachten positief of negatief waren, bleek veel belangrijker te zijn.

Alles bij elkaar genomen, is de bezorgdheid over de negatieve effecten van online waakzaamheid wellicht niet gerechtvaardigd. Online waakzaamheid heeft over het algemeen een negatief effect op het welzijn, maar dit betreft een indirect en klein effect. In het dagelijks leven is online waakzaamheid in de vorm van gedrag niet problematisch gebleken. Online waakzaamheid in de vorm van gedachten is op negatieve wijze aan welzijn gerelateerd, maar dit effect was klein en minder belangrijk dan de kwaliteit van deze gedachten. Positieve gedachten over onlinecommunicatie blijken veel belangrijker dan de frequentie waarin deze voorkomen.

DOEL 2: HEBBEN SMARTPHONESIGNALLEN INVLOED OP PRESTATIES?

In hoofdstukken 4 en 5 hebben we getoetst of smartphonesignalen (meldingen en de aanwezigheid van een smartphone) de cognitieve prestaties schaden. De onderzoeken toonden een verrassende discrepantie aan tussen ervaring en prestaties. In hoofdstuk 4 rapporteerden deelnemers een veel hogere mate van online waakzaamheid en afleiding wanneer hun telefoon een melding ontving of zelfs wanneer hun telefoon alleen op tafel lag dan de controlegroep zonder telefoon. Hun prestaties werden echter niet beïnvloed door hun telefoons. Hoofdstuk 5 laat een vergelijkbaar resultaat zien: In tegenstelling tot deelnemers uit de controlegroep, ervaren proefpersonen uit de experimentele groep de meldingen als zeer afleidend, hoewel hun prestaties daar niet onder leden.

Alles bij elkaar genomen, lijken smartphonesignalen niet problematisch voor onze prestaties. Hoewel deelnemers het gevoel hebben dat de aanwezigheid van hun smartphones of meldingen afleidend zijn en een verhoogde alertheid tot gevolg heeft, hebben dergelijke signalen geen nadelige invloed op de prestaties.

DOEL 3: ZIJN SMARTPHONESIGNALLEN BELONEND?

In hoofdstuk 6 hebben we getoetst of smartphonesignalen belonend zijn, wat zou kunnen verklaren waarom mensen meldingen en hun telefoons als storend ervaren. Onze resultaten laten ook hier een discrepantie zien. Mensen beoordeelden sociale smartphone-apps (zoals Facebook, Instagram) als veel meer belonend dan neutrale smartphone-apps (zoals de rekenmachine, klok), vooral wanneer deze sociale smartphone-apps meldingen toonden. Dit verschil in zelfgerapporteerde beloning werd echter niet omgezet in gedrag. Hoewel sociale smartphone-apps als zeer belonend werden beschouwd, trokken deze apps niet meer aandacht dan neutrale apps of apps die minder belonend werden gevonden.

Er kan worden geconcludeerd dat hoofdstukken 4 en 5 een discrepantie tussen zelfgerapporteerde afleiding en prestaties lieten zien, en uit hoofdstuk 6 een discrepantie is gebleken tussen beloningsclassificaties en het falen van die beloningen om de aandacht

te trekken. Smartphonesignalen kunnen als belonend worden ervaren, maar zijn niet belonend genoeg om ons af te leiden.

DOEL 4: KUNNEN WE VORKEUREN VOOR SMARTPHONESIGNALEN REDUCEREN?

In hoofdstuk 7 hebben we onderzocht of een zogenaamde go-/no-go-training de aantrekkingskracht van smartphonesignalen kan verminderen. We hebben specifiek onderzocht of deze training de expliciete zelfgerapporteerde voorkeur voor smartphone-apps kan verminderen. Na het uitvoeren van twee experimenten ontdekten we dat de training effectief was in het reduceren van de mate waarin proefpersonen bepaalde apps leuk vonden. Deze voorkeur was belangrijk, omdat proefpersonen minder gebruikmaakten van apps die ze minder leuk vonden. Met andere woorden, de training verminderde de mate waarin proefpersonen bepaalde apps leuk vonden, wat tot gevolg had dat proefpersonen deze verminderde voorkeur als leidraad hanteerden in de keuze van de te gebruiken app.

Samengevat suggereert hoofdstuk 7 dat de go-/no-go-training effectief is in het verminderen van de zelfgerapporteerde aantrekkingskracht van smartphonesignalen. Deze vermindering kan belangrijk zijn; deze kan mogelijk de waargenomen afleiding van smartphonesignalen verminderen en zelfs beïnvloeden op welke smartphonesignalen mensen besluiten te reageren.

CONCLUSIE

Elke nieuwe technologie die snel aan populariteit wint, brengt enige bezorgdheid met zich mee. Deze bezorgdheid is er ook voor smartphones. Onze resultaten suggereren dat bezorgdheden rondom smartphonesignalen en online waakzaamheid wellicht niet gerechtvaardigd zijn. Op basis van de bevindingen uit dit proefschrift zou het voorbarig zijn om drastische maatregelen te treffen, zoals het verbieden van telefoons in openbare ruimtes of op scholen. Buiten het laboratorium lijkt online waakzaamheid geen problemen voor het welzijn op te leveren. In het laboratorium lijken mensen potentiële afleidingen van smartphones te kunnen negeren.

Summary

Smartphones are everywhere and many people are concerned that they might be detrimental to our performance and well-being. *Smartphone cues*, such as notifications or even the mere presence of a phone, have shown to be distracting. Moreover, these cues can lead to a state of constant alertness, labeled *online vigilance*. Users experience both smartphone cues and the resulting online vigilance as bothersome. However, there is little research testing whether concerns about negative effects on performance and well-being are warranted. Therefore, this dissertation had the overall aim to investigate the effects of smartphone cues and online vigilance on well-being and performance.

To address this question, we had four goals. First, we investigated the relation of online vigilance with well-being on the trait (Chapter 2) and on the state level (Chapter 3). Second, we tested whether smartphone cues interfere with the executive control functions of inhibition (Chapter 4) and working memory (Chapter 5). Third, we examined whether online vigilance can be explained by the rewarding nature of smartphones (Chapter 6). Fourth, we tested whether we can reduce people's preferences for smartphone cues (Chapter 7).

GOAL 1: DOES ONLINE VIGILANCE RELATE TO WELL-BEING?

In Chapter 2, we tested whether online vigilance would be related negatively to well-being if it takes the form of absentmindedness, specifically mind-wandering and mindfulness. To obtain a more comprehensive picture of possible relations with well-being, we assessed both how participants evaluate their lives (i.e., cognitive component) and how participants usually feel (i.e., affective component). As predicted, online vigilance and the two well-being components were related indirectly: Online vigilance was related to the well-being outcomes through decreased mindfulness, but not through mind-wandering.

Chapter 3 relied on a combination of in-the-moment self-reports and objective smartphone use to assess the relation between online vigilance and well-being in everyday life. Our results show that online vigilance, when expressed as a behavior such as checking your phone or responding to notifications, was not related to well-being. When expressed in thoughts about online communication, online vigilance was negatively related to well-being. The more people thought about online interactions in the past half an hour, the worse they felt in the current moment. However, this relation was small and it was far more important whether such thoughts were positive or negative.

Taken together, concerns about negative effects of online vigilance might not be justified. When considering online vigilance in general, it relates negatively to well-being, but only indirectly and to a small extent. In everyday life, online vigilance expressed as behavior was not problematic. Expressed as thoughts, online vigilance was negatively related to well-being, but this effect was small and less important than the quality of these thoughts. Positive thoughts about online communication are far more important than how frequently they occur.

GOAL 2: DO SMARTPHONE CUES IMPACT PERFORMANCE?

In Chapters 4 and 5, we tested whether smartphone cues (notifications and the presence of a smartphone) would impair cognitive performance. The studies showed a surprising discrepancy between experience and performance. In Chapter 4, participants reported much higher online vigilance and distraction when their phone received a notification or even when their phone was only on the table compared to a group without their phones. However, their performance was not affected by their phones. Chapter 5 showed a similar result: Participants perceived notifications as highly distracting compared to a control group, but their performance did not suffer.

Taken together, smartphone cues do not appear problematic for our performance. Although participants feel that the presence of their smartphones or notifications distract them and put them in a state of heightened alertness, such cues did not impair performance.

GOAL 3: ARE SMARTPHONE CUES REWARDING?

In Chapter 6, we tested whether smartphone cues are rewarding, which could explain why people perceive notifications and their phones as distracting. Our results show yet another discrepancy. People rated social smartphone apps (e.g., Facebook, Instagram) as much more rewarding than neutral smartphone apps (e.g., Calculator, Clock), particularly when these social smartphone apps had a notification sign. However, this difference in self-reported reward did not translate to behavior. Although rated as very rewarding, social smartphone apps did not attract more attention than apps rated as less rewarding or neutral apps.

Taken together, whereas Chapters 4 and 5 showed a discrepancy between self-reported distraction and performance, Chapter 6 showed a discrepancy between reward ratings and the failure of those rewards to capture attention. Smartphone cues might be perceived as rewarding, but not as rewarding enough to distract us.

GOAL 4: CAN WE REDUCE PREFERENCES FOR SMARTPHONE CUES?

In Chapter 7, we tested whether a so-called go/no-go training can reduce the appeal of smartphone cues. Specifically, we investigated whether the training can reduce explicit self-reported liking of smartphone apps. In two experiments, we found that the training was effective in reducing how much participants liked certain apps. This liking was

important because participants also chose to use these apps less if they liked them less. In other words, the training reduced how much participants liked certain apps; participants then relied on this liking to make a choice of what app to use.

Taken together, Chapter 7 suggests that the go/no-go training is effective in decreasing the self-reported appeal of smartphone cues. This reduction may be important; it has the promise to reduce the perceived distraction of smartphone cues and could even help influencing what smartphone cues people choose to engage with.

CONCLUSION

There have been concerns surrounding every new technology that rapidly gained popularity. These concerns also surround smartphones. Our results suggest that worries around smartphone cues and online vigilance may not be justified. Based on the findings in this dissertation, it would be premature to implement drastic measures, such as banning phones in public spaces or schools. Outside the laboratory, online vigilance does not seem problematic for well-being. In the lab, people seem capable of ignoring potential smartphone distractions.

Acknowledgements

Writing a dissertation by yourself is possible, but not worth it. That's why I'm grateful that I had help along the way. Not being a man of many (kind) words, I'll try to make this short and entertaining – unlike my dissertation.

I am grateful to have had two great supervisors. Moniek, thank you for trusting me to do my thing and for intervening when that thing wasn't good. Without your support, this thesis would be much worse. Harm, you're an extremely cool dude, and I'll really miss working with you. I hope we can find an arrangement for you to be my life coach from now on.

I am equally grateful to have had the very best of office mates. Thabby, Faart, they say a job is only as good as the people you work with. I agree. We shall be grumpy until the end of time or until we die, whichever comes first.

I should not forget to thank my Communication colleagues. Communication is a weird, but fun group, and I am lucky to have been part of that group. D, J, in my first year we spent more time together outside office hours than during office hours. I'm glad we did. Danielle, Rhianne, thanks for bringing some desperately needed sanity to our PhD group. Sandy, I don't know you very well, but you seem nice. And of course, Susanne: Thank you for your patience and help over the past four years.

The BSI is a large institute, which means I made friends outside Communication. Mikey, you might've left science, but you'll never leave my nightmares. Nor will the incriminating evidence ever leave my phone. Papa Bear, your magnificent biceps are only matched by your good looks. Nicole, your biceps are less impressive, but feel free to write yourself something nice:

Yoni, once you become a male model known only as The Shredded Weasel, I expect royalties for the name. Meta, thank you for always being there. And by there, I mean both your office and any occasion for drinks. Ronny, you might have called me Johannes for the full four years, but you made time in the lab much better.

Dori, you're the best thing that came out of those four years.

Mein Dank geht selbstverständlich auch an Freunde, Familie und jene, die in keine dieser Kategorien fallen, aber dennoch wichtig sind. Die Mannheim-Amsterdam-Gang fällt definitiv in die erste Kategorie. Hannah, danke für Kuchen, wenn ich ihn brauchte, und für Kritik, wenn ich sie nötig hatte. Märker, danke für viele Biere von niedriger Qualität und Freundschaft von hoher. Verena, Nadine und Schacti, das gleiche gilt für euch, nur mit gutem Wein statt schlechtem Bier. Andre, danke für die letzten 20 Jahre. Ich freue mich auf die nächsten 20, in denen ich dich auf der Xbox abziehe. Lude, vergiss den Controller nicht. Herr Vorderer, Sie haben mir die ersten Schritte in der Wissenschaft ermöglicht – danke dafür.

Und die Familie. Mama, ich hoffe, jetzt bin ich endlich erfolgreich genug für deine Zuneigung. Danke, dass du immer für mich da warst. Papa, ich habe von dir das volle Haar und du von mir einen kaputten Rücken wegen zahlloser Umzüge. Danke für deine Hilfe, wann immer ich sie brauchte. Irina, allein wäre es bis hierher langweilig gewesen. Ich bin froh, dass du da warst – und bist. Fredi, wenn am Ende des Monats das Telefon klingelte,

hattest du immer Rat und genauso oft deine Kreditkarte parat. Ersterer war wichtiger. Vielen Dank. Alex, Uli, Richie, Jürgen, Ari, in den letzten vier Jahren brauchte ich vor allem Alkohol, Futter und ein offenes Ohr. Danke, dass ihr mich mit allen dreien versorgt habt.

About the Author

I was born on the 21st October 1989 in Mainz, Germany. In 2013, I obtained my bachelor's degree in Media and Communication and German Literature from the University of Mannheim. I obtained my master's degree in Communication from the University of Amsterdam in 2015. From 2015 until 2019 I was working on obtaining a PhD in Behavioral Science from Behavioural Science Institute, Radboud University, the results of which you hold in your hands.

Behavioural
Science
Institute

Radboud University

