Playing smartphone games while studying: an experimental study on reading interruptions by a smartphone game

Katharina Graben¹ · Bettina K. Doering² · Antonia Barke²

Received: 21 July 2021 / Accepted: 26 September 2021 / Published online: 11 October 2021
© The Author(s) 2021

Abstract

In this study, we investigated whether the use of smartphone games while reading a text reduces learning performance or reading speed. We also examined whether this is affected by push notifications. Ninety-three students were randomly assigned to three learning conditions. In the gaming group (G), participants played a game app for 20 s at 2-min intervals while reading. In one subgroup, the game app sent push notifications (GN+); in the other subgroup, no notifications (GN–) were sent. In the control group (C), participants did not play a game. After the reading, participants took a multiple-choice quiz. We compared quiz scores and reading times of the groups (G) and (C) and within the gaming group (GN+, GN–) and observed no differences. Since the statistical non-significance of these tests does not entail the absence of an effect, we conducted equivalence tests, which did not demonstrate equivalence either. The experiment ensured high internal validity, yet remained inconclusive. Reasons for the similarity of performance in all groups could be non-specific exercise effects (all participants owned a smartphone), low similarity between the tasks, low variance of participants' ability and motivation (high achieving, low ADHD scores) or low game complexity. Future research should address these questions.

Keywords Smartphone · Multi-tasking · Lecture · Interruption · Mobile phone games

✉ Katharina Graben
katharina.graben@uni-marburg.de

¹ Division for Clinical Psychology and Psychotherapy, Faculty of Psychology, Philipps-University Marburg, 35032 Marburg, Germany

² Clinical Psychology, Catholic University Eichstaett-Ingolstadt, Eichstaett, Germany

1 Introduction

While reading this study, where is your smartphone right now? Are you sure that it will not interrupt your reading? Smartphones have become nearly ubiquitous, especially in the life of younger persons. If you visit a university library these days, the chances are high that you will see students studying while their smartphones lie next to the book, taking a look at their phones at more or less regular intervals. According to one study, students interrupt their learning due to modern technologies like smartphones or computers every 6 min (Rosen et al., 2013). This is considered a problem, because interruptions may impair learning or memory performance (Conard & Marsh, 2014; Kuznekoff & Titsworth, 2013; Oulasvirta & Saariluoma, 2004). Nearly all German households (96.7%) own at least one mobile phone (Destatis, 2018) and especially in the age group of 14–49 year olds, smartphones (over 95% users) have become indispensable (VuMa, 2019). Considering these numbers, the question if smartphones interfere with learning is highly relevant.

1.1 Interruptions due to smartphones

The effects of distracting interruptions on performance in a main task have been investigated from several perspectives. McFarlane (1999) examined different types and modes of interruptions: ‘immediate’ (requiring an immediate response from the user), ‘negotiated’ (the users decide for themselves when they want to react), ‘mediated’ (an intelligent system decides a favourable time for the interruption) and ‘scheduled’ (regular interruptions after specified times). The study found that all types of interruptions had negative effects: depending on the type of interruption, people made more mistakes in one of the tasks (the interrupted one or the interrupting one) or they needed more time for the task completion. In all forms, however, the interruption seemed to interfere with performance. Another study also found that interruptions may slow down participants working on a main task (Kreifeldt & McCarthy, 1981). A few studies have investigated interruptions by smartphones, specifically in learning contexts. Many rely on self-report data, using correlational designs. In a correlational study with about 2000 US college students, the college grades and self-reported information regarding their use of Facebook, emailing, talking on their smartphones and texting while doing college homework were recorded (Junco & Cotten, 2012). The use of Facebook and texting while doing homework was associated with worse grades. Further self-report studies also found a negative relationship between the use of social media and the grade point average (Karpinski et al., 2013; Rosen, et al., 2013). Turning to experimental evidence, Wood et al. (2012) compared the performance after a 20-min lecture presentation on research in a 15-items quiz between an experimental group using a smartphone (texting, emailing, MSN messaging and Facebook) with three control groups (enforced paper-and-pencil note-taking, enforced word-processing note-taking and a natural use of technology condition). The control groups outperformed the smartphone group, indicating that mobile phone use may be detrimental to learning performance. These results have been replicated with similar experimental designs

(Conard & Marsh, 2014; Dietz & Henrich, 2014; Gupta & Irwin, 2016). Czerwinski et al. (2000) used a repeated-measures design in which participants working on a computer task (searching for book titles in a spreadsheet) were interrupted by text messages from computer searches (interruption condition) or not interrupted (control condition). Interruptions were associated with a decline in the participants' task performance (searches with interruptions took significantly longer than searches without interruptions). Of note, all these studies restricted smartphone use to communication (texting, using social media). However, other frequent smartphone uses include mobile phone games (Ipsos, 2016): Every fourth German Internet user plays daily video games on a smartphone and the majority (64%) of German app players are between 18 and 44 years old.

Games may lead to frequent interruptions due to a game-inherent mechanism: in many gaming apps, a point mechanism is implemented, in which a kind of 'energy' is expended through game play. When depleted, it will recharge itself over time. Once recharged, the user can continue to play (for example *Harry Potter: Hogwarts Mystery*TM requires an energy unit every 4 min). In addition, it may take time to collect a reward (for example in *Hay Day*, planted wheat takes 2 min and carrots 10 min before they can be harvested). Many gaming apps signal these events by a so-called 'push notification' intended to turn the user to the game again. Such notifications can entail audio or vibration signals, depending upon the user's phone settings, and may thus become so-called external interruptions to other activities. The notifications can be turned off or on manually in the user-settings of the mobile phone or the respective app. Even if users switch off the notifications sent by the game, however, they may still be aware of the respective intervals and check into the game accordingly in order to advance in the game. In these cases, the interruptions may become internally generated self-interruptions.

Interruptions are generally detrimental to the interrupted task (Couffe & Michael, 2017; Trafton & Monk, 2007). However, the 'costs' to the main task differ between internally and externally generated interruptions. Katidioti et al. (2016) found a longer processing time of the main task with internally controlled interruptions in contrast to external interruptions. The authors explain the longer processing time by the higher cognitive workload that is required to make the conscious decision to change tasks. However, contrary results have also been reported. In a natural office environment, participants needed more time to return to the main task after the interruption, when the interruption was external and not internal. Possibly, the participants had more difficulty returning to their actual task if they were not in control of the time of the interruption (Cades et al., 2010).

1.2 Study goals

Regarding game apps, it is unknown whether a casual gaming app causes interference with study tasks (and if so, to what extent) and how receiving push notifications compares to interrupting the main task on one's own initiative to check the game. In the present study, we examined the effects of a standardized smartphone game on students' performance when studying a text. In addition, we investigated

whether the use of app notifications resulted in performance differences. Performance was operationalized as learning performance (i.e. scores in a quiz testing the participants' understanding and retention of information contained in the studied text) and reading time.

We tested the following hypotheses with regard to the differences between students using a smartphone game (G) and students in the control group (C), who did not use a smartphone game:

1. Students' learning performance will be lower when they use the smartphone game.
2. Students' reading time for the text will be longer when they use the smartphone game.

We also tested whether the subgroups (GN+) and (GN–) differed in performance or reading time. Due to the inconsistent evidence regarding the consequences of the locus of the interruptions, we tested two-sided hypotheses.

3. Students' learning performance will differ depending on whether they use the smartphone game with or without notifications.
4. Students' reading time for the text will differ depending on whether they use the smartphone game with or without notifications.

2 Method

2.1 Procedure and participants

The study was conducted in a behavioural laboratory at Philipps University Marburg. For recruitment of a student sample, the study was advertised online on the university Research Participation System and via a university-wide email list. After providing informed consent, the participants were able to familiarize themselves with the game app by playing a demo version of the game (description below) on a provided smartphone. The questionnaires were implemented in the online survey software *SoSci Survey* (SoSci Survey GmbH, Munich, <https://www.soscisurvey.de>) and filled in at a computer. The participants provided demographic information and rated their prior knowledge regarding the topic of the text (i.e. methane hydrate). The participants were then asked to read a text on methane hydrate. All participants were informed that their learning performance would be tested after the reading and that they should indicate to the experimenter when they had finished reading. Participants allocated to the GN+ and GN– groups read the text while playing the smartphone game every 2 min for 20 s. They received the instruction that both tasks (reading and gaming) were important and they should try to do as well as possible in both of them. Participants of the control group (C) did not play the game. Subsequently, all participants answered a quiz about the text. After completing each experimental part (reading plus gaming and quiz), they provided information about their motivation and their subjective performance, as well as their typical smartphone use. Finally, the participants filled in a screening test for Attention Deficit Hyperactivity Disorder for use as a control variable (see below).

Inclusion criteria were age over 18 years, being student at a German university, enrolled in German-speaking courses and unrestricted abilities of seeing, hearing and movement of the fingers. The participants were allocated randomly to one of the three experimental groups (notifications GN+, no notifications GN–, control C). A total of 98 participants provided informed consent; three participants had to be excluded because of technical difficulties and two because the participants did not follow the instructions. Thus, 93 participants remained to be analysed, 31 per group (GN+, GN–, C). The participants were aged 22.8 ± 3.8 years; the majority was female (73.1%). The mean school-leaving grade of the participants was 1.8 ± 0.7 (possible range: 0.7 [best] to 4.0 [worst]). All participants reported owning a smart-phone (although this was not an inclusion criterion).

2.2 Material

2.2.1 Demographic information and prior knowledge

The participants provided information on sex, age, first language, subject studied, semester and Abitur school leaving grade (Abitur: the German school leaving certificate, entitling the person to study at a university). In order to assess prior knowledge of the topic of the text (i.e., methane hydrate), we asked participants to rate their knowledge about methane hydrate and six distractors (cinema, the Tau-Ceti system, fracking, low-energy houses, neopterans, forest owls) on a six-point scale (0 = ‘I know nothing about this topic’, 5 = ‘I know everything about this topic’).

2.2.2 Text

As study text, we selected a text that was similar in complexity to texts students were used to in study courses. The topic was selected with the aim to be unfamiliar to most participants to ensure that any knowledge on the topic had to be gained from the text. To choose the text, we conducted a pilot study, in which 15 students of psychology at the Philipps University Marburg were asked to indicate their knowledge of three topics on a six-point rating scale (0–5). For ‘methane hydrate’, the pilot test participants indicated virtually no prior knowledge (0.13 ± 0.34). Based on this result, we decided to use a shortened version (1722 words) of the article ‘Brennendes Eis: Methanhydrat—Energiequelle der Zukunft oder Gefahr fürs Klima?’—‘Burning ice: methane hydrate—energy source of the future or danger for the climate?’ (Gutt et al., 2001).

2.2.3 Quiz

A knowledge test in multiple-choice format tested the participants’ learning performance with regard to the text. Participants had to choose the correct answer from among four answer options for each of 18 questions. The learning performance was indicated by quiz scores, i.e., the number of correctly answered questions (maximum 18, minimum 0).

2.2.4 Smartphone game

In order to ensure a standardized gaming app novel to all participants, the first author (KG) programmed an easy custom gaming app (see Fig. 1) using the MIT App Inventor (Massachusetts Institute of Technology, Massachusetts, USA). The app was installed on a smartphone (Huawei P8 lite 2017), which was provided to all participants. The participants' task was to drag a penguin figure across the screen with a finger to 'pick up' statically displayed fish while avoiding a moving polar bear. The number of fish collected appeared as a score at the top of the screen. Whenever the polar bear was touched, the score was reset to nil. The game was playable for 20 s until a black screen appeared for at least 2 min. Participants could return to the game screen by simply tapping the screen when the 2-min wait was over. Depending on the test conditions, the end of the waiting time was either signalled by a short vibration and a beep (GN+), or not announced (GN–).

Fig. 1 Screenshots of the smartphone game. Left panel: waiting screen. Right panel: gaming screen. 'Gesammelte Fische': number of fish collected; 'fertig': finished

2.2.5 Motivation check

After each part of the experiment (i.e., the text, the quiz and the last gaming episode), participants reported how motivated they were in completing the preceding experimental activity, how much fun the activity was, how interested they were in the respective activity. After the quiz, they also rated how well they believed to have done on the quiz. The participants in the gaming group also indicated how much they felt disturbed by the app while studying and how well they believed to have done in the game. All questions were rated on a six point scale (0 = ‘not at all’, 5 = ‘very much’).

2.2.6 Smartphone use

Participants completed the *Problematic Use of Mobile Phones* (PUMP) scale (Merlo, et al., 2013) to assess problematic user patterns. The scale consists of 20 statements about possible thoughts, feelings, and behaviours related to problematic smart phone use, which have to be rated on a five-point scale, from 1 = ‘strongly disagree’ to 5 = ‘strongly agree’. The PUMP scale demonstrated good internal consistency, with $\alpha = .94$. In the present study, we used the German translation PUMP-D (Graben et al., 2020), which also showed good psychometric properties, with $\omega = .91$.

In addition, the participants were asked whether they owned a smartphone (mobile phone with internet access), a cell phone (mobile phone without internet access) or no mobile phone at all. Moreover, participants reported how much they used their mobile phone on a typical day (in minutes), how often they typically interrupted their studies due to self-initiated interruptions (count) and how often interruptions occurred as a reaction to signals or notifications from their mobile phone (count).

2.2.7 Attention Deficit Hyperactivity Disorder (ADHD)

To screen for attention deficits, we used the German 30-item screening version of the Conners Adult ADHD Rating Scales (CAARS) (German scale: Christiansen et al., 2011; original scale: (Conners et al., 1999). The frequency and severity of any ADHD-related symptoms are rated on a four-point scale (from 1 ‘not at all/ never’ to 4 ‘very much/ very frequently’). The self-report questions can be organized in five different scores (*DSM-IV* ADHD Symptoms score, Inattention score, Hyperactivity/Impulsivity score, Index score) and a total symptom score. The scores have good psychometric properties with Cronbach’s alpha between .74 and .95 (Adler et al., 2008) and excellent test–retest reliability with $r = .89$ (Erhardt et al., 1999).

2.3 Data analysis

To adjust the text reading time for the time spent playing the game, we subtracted the playing time (20 s per round) from the total reading time for each participant and, in the following, always referring to ‘adjusted reading time’.

The groups were compared for baseline characteristics and successful randomization regarding basic variables with one-way ANOVAs regarding age, Abitur school

leaving grade, ADHD score, previous knowledge and interest in the topic of the text, and the motivation to perform well in the quiz and with the χ^2 test (sex). The one-sided hypotheses (1) and (2) regarding gaming vs. non-gaming were tested by comparing group G with Group C by one-tailed t -tests; the two-sided hypotheses (3) and (4) by comparing the subgroups GN+ and GN– with two-tailed t -tests. As a measure of effect size, Cohen's d is reported.

To investigate the absence of an effect of the experimental manipulation, the analyses were repeated with group-specific and overall outliers ($\pm 2 \times \text{SD}$) excluded. Additionally, post hoc equivalence tests comparing group G with Group C and comparing the subgroups GN+ and GN– with the dependent variables adjusted reading time and quiz scores were computed (Lakens, 2017).

Missing data were excluded on a pairwise basis. All analyses were computed with SPSS version 21.0.0 (IBM, Meadville, USA).

3 Results

The experimental groups did not differ with regard to age, Abitur school leaving grade, sex, or any of the other control variables (see Table 1 for a full characterization).

3.1 Quiz performance and adjusted reading time for the textual task

The quiz scores of the control group C (11.90 ± 2.67) and the gaming group G (11.44 ± 2.77) did not differ [one-tailed t -test: $t(91) = 0.78$, $p = .219$, $d = 0.17$]. The adjusted reading time was descriptively, but not statistically significant longer for the gaming group (13.1 ± 4.0 min) than for the control group (12.2 ± 4.0 min) [one-tailed t -test: $t(91) = -1.14$, $p = .129$, $d = -0.27$]—for z -values see Fig. 2. Participants' quiz scores did not differ between those who received notifications (GN+: 11.23 ± 2.80) and those who did not (GN–: 11.65 ± 2.76) [$t(60) = 0.59$, $p = .555$, $d = 0.15$]. Likewise, the adjusted reading times did not differ between GN+ (13.0 ± 4.3 min) and GN– (13.3 ± 3.7 min) [$t(60) = 0.30$, $p = .764$, $d = 0.08$]—for z -values see Fig. 3.

3.2 Performance regarding the gaming app

Only participants of the GN+ and GN– group used the gaming app. In every round, a maximum of 43 fish could be collected. Participants collected on average 34.4 ± 29.8 fish per round (high variance due to the possibility to “hit the polar bear” at the end of a round what resulted in no collected fishes) and played a mean of 6.4 ± 1.3 rounds. On average, they collided with the polar bear once (± 1.3) in the whole time, which corresponds to 0.15 collisions per round. In none of these measures were significant group differences between GN+ and GN– observed [fish per round: $t(60) = 1.06$, $p = .294$, $d = 0.27$; number of played rounds: $t(60) = -0.95$, $p = .346$, $d = 0.24$; collisions with the polar bear per round: $t(60) = -0.54$, $p = .592$, $d = 0.14$]. The participants in the GN– group took more time to begin playing (first

Table 1 Characterization of control variables

	C		GN–		GN+		F(2,90)	p	η^2
	M	SD	M	SD	M	SD			
Age	23.13	4.86	21.42	2.53	22.84	3.56	1.823	.168	0.039
Abitur school leaving grade	1.94	0.67	1.80	0.69	1.74	0.73	0.657	.521	0.015
ADHD score	21.26	10.02	24.16	12.30	23.35	10.62	0.573	.566	0.013
Previous knowledge of text topic	0.19	0.48	0.35	0.76	0.16	0.37	0.067	.349	0.023
Interest in the text	2.19	1.32	1.90	1.25	2.13	1.46	0.398	.673	0.009
Motivation regarding the quiz	3.68	1.19	3.65	1.23	3.35	1.20	0.672	.513	0.015
Usage of smartphone per day (min)	136.77	91.693	173.06	81.60	163.71	74.34	1.603	.207	0.034
Number of interruptions per day without notifications	7.10	6.00	10.45	9.24	13.77	12.95	3.586	.032	0.074
Number of interruptions per day because of notifications	8.97	11.67	14.81	18.94	9.06	9.76	1.761	.178	0.038
Preference for distraction-free learning environment	3.65	1.02	4.03	1.08	4.03	0.80	1.639	.200	0.035

Characterization of control variables with their mean and standard deviation per group, F value, p value and η^2

Fig. 2 Adjusted reading time and quiz performance (G/C). Adjusted reading time and quiz performance of the gaming group (G) and the non-gaming control group (C). Z-values with standard errors shown

Fig. 3 Adjusted reading time and quiz performance (GN-/GN+). Adjusted reading time and quiz performance of the two gaming groups without (GN-) and with push notifications (GN+). Z-values with standard errors shown

touch of the screen) after the game became playable again $t(60)=5.96$, $p<.001$, $d=1.68$. Both groups were equally motivated to play the game ($t(60)=-0.667$, $p=.507$, $d=0.17$) with a mean of 4.78.

On a scale from ‘not at all’ (1) to ‘very much’ (6), participants felt 4.6 ± 1.5 distracted by the gaming-app while reading the text. To identify whether participants felt more distracted by the notifications (externally controlled disruption) or by monitoring the time (internally controlled disruption), a *t*-test was conducted between the two gaming groups (GN+, GN–), but no difference was observed [$t(60) = -1.35$, $p = .183$, $d = -0.35$]. How much the participants felt distracted by the game correlated with adjusted reading time ($r = .40$, $p = .001$) but not with the quiz-score ($r = -.05$, $p = .706$).

3.3 Post-hoc analyses

We repeated the analyses with group-specific and overall outliers ($\pm 2 \times SD$) excluded, which did not change of any of the results.

Given the absence of significant group effects, we conducted equivalence tests based on participants’ mean quiz scores and adjusted reading time across the groups (Lakens, 2017). The test was conducted with Welch’s *t*-test (corrected for unequal variances). The smallest effect size of interest (SESOI) for the quiz score was regarded as a difference of one point in the quiz score, resulting in the equivalence bounds $\Delta_L = -1$; $\Delta_U = 1$. The SESOI for adjusted reading time was determined as Cohen’s $d = 0.2$ and resulted in the following equivalence bounds: For the comparison of groups C and G $\Delta_L = -42.57$ and $\Delta_U = 42.57$; for the group comparison GN+ and GN– $\Delta_L = -47.99$ and $\Delta_U = 47.99$. For the quiz scores, the equivalence tests were neither significant for the comparison of groups C and G ($t(62.23) = -0.90$, $p = .187$), nor for the comparison of groups GN+ and GN– ($t(59.99) = -0.82$, $p = .207$). For the adjusted reading time, the equivalence tests were also not significant (comparison of groups C and G: $t(86.39) = -0.27$, $p = .395$; comparison of groups GN+ and GN–: $t(58.65) = -0.49$, $p = .315$).

4 Discussion

With this study, we present the first experiment to investigate whether the parallel use of smartphone gaming apps disturbs students’ learning performance when reading a text. We did not find a reduced reading speed for gaming app users or a decrease in their test performance compared to a control group. Moreover, the gaming application’s notification mode did not exert an influence on reading time or quiz performance.

With regard to the existing research (Conard & Marsh, 2014; Karpinski et al., 2013; Kuznekoff & Titsworth, 2013; Oulasvirta & Saariluoma, 2004; Rosen et al., 2013), it is unexpected that no significant difference was found between the group that could read the text without distraction and the group playing on the smartphone while reading. It is noteworthy that the quiz scores and reading times cannot be declared as equal (see equivalence tests), either.

We therefore examined design-specific or sample-specific characteristics as possible reasons for this unexpected finding. There is a possibility that the distracting

effect from the parallel use of smartphone games while reading is actually present, but smaller than previously assumed and thus not significant in our relatively small sample. Alternatively, one may speculate that the distractive effect of the gaming app may have been present, but that our student sample (consisting of participants highly motivated to perform well in achievement tasks) was able to compensate for it. An additional finding in our data strengthens this interpretation: participants' ratings of how much they felt disturbed by the app correlated moderately ($r = .40$) with reading time. This could be indicative of the aforementioned compensation of a disruptive effect, resulting in increased cognitive load: people tend to drive more of their cognitive resources into tasks when challenge stressors appear, which helps with short term performance but produces faster mental fatigue (Crawford et al., 2010; Widmer et al., 2012). Mental fatigue is already known to reduce working memory performance (Borragán et al., 2017; Faber et al., 2012) and could possibly result in poorer quiz scores and/or longer reading time in the case of a longer text (since this increased effort could not be maintained in the long term). Thus, one might speculate that less-able or less motivated learners may not compensate as successfully. In this respect, our high-achieving sample will probably underestimate the effect of the parallel use of gaming apps on learning performance.

In addition, the participants of our study may already have been somewhat used to such interruptions, as all participants owned a smartphone and evidence suggests that training reduces interruptive effects. Hess and Detweiler (1994) showed that, after two training sessions with interruptions, the interruptions lost some of their detrimental effect. In the current study, only participants owning a smartphone participated, and the study design was close to a natural learning situation. It is highly likely that participants were somewhat accustomed to interruptions by their smartphone while reading texts and had experience in the compensation required.

Another reason for the unexpected result of not finding a direct effect on learning performance may be that, in previous experiments focussing on reading/learning tasks, the interruptions were linked to texting (Conard & Marsh, 2014; Dietz & Henrich, 2014; Gupta & Irwin, 2016; Wood et al., 2012) and, in their cognitive demands, were thus very similar to the main task. Playing a simple game may be less disruptive than texting in these contexts because reading a text (as main task) and playing a smartphone game (as a disruptive task) are less similar. The more similar the interruptive task is to the main task, the more disruptive it seems to be (Gillie & Broadbent, 1989).

To elucidate whether the push notifications of games influence the distractive effect, we experimentally manipulated these app settings to investigate whether they resulted in performance differences. However, no difference emerged between the groups with and without notifications, indicating that both push notifications (i.e. externally controlled disruptions) and the participant's own time monitoring (i.e. internally controlled disruptions), seemed equally (non-)disruptive. Possibly, receiving a push notification is a clear interruption, whereas in the GN- condition, in which participants had to monitor when the game was playable again, one could argue whether it is an internally generated interruption or more akin to a dual-task. However, in this context it is also interesting to consider that quiz scores and reading time between the two gaming groups cannot be declared

to be the same, either. Thus, we cannot be sure that there is no detrimental effect of push notifications. Given that it is unlikely that smartphones will disappear from lecture halls, classrooms and work desks in the near future and that a detrimental effect cannot be excluded, our study points to a variable that can be easily changed by the user (push notifications off or on). For future research, it is important to investigate this type of variable (also for example interruption by vibration or by sound) and its effects, as people could potentially be more willing to change a setting on their smartphone than to ban it completely during studying.

In the current study, factors that might compromise internal validity were strictly controlled: all participants used the same experimental smartphone and the same gaming app. No participant had prior experience regarding the gaming app or any knowledge regarding the topic of the text and the randomization resulted in groups comparable with regard to important basis variables. A few limitations should be taken into account when interpreting our results. Firstly, for reasons of standardization, the participants did not use their own smartphones. This may have minimized the size of the effect because evidence suggests that many people feel a strong psychological attachment or involvement towards their personal mobile phone (Fullwood et al., 2017; Walsh et al., 2010) and therefore may pay more attention to their own smartphone. Secondly, most of the participants were psychology students (58.1%) with very good Abitur school leaving grades, who may not be representative of the general population of university students. It is possible that these academically very able and highly motivated students are particularly capable of compensating for the negative effects of mobile phone use. In particular, they showed good impulse control and unremarkable ADHD scores. Thirdly, the game design should be mentioned: it was a very simple game that might only use a few cognitive resources. A recent review (King & Delfabbro, 2014) found that a higher investment of time, energy or money in a (Internet) game is one of the factors leading to increased playing, which is already known as the sunk cost effect (Arkes & Blumer, 1985). In our study, the game was new to all participants (i.e. no previous investment of time, energy or money had taken place) and, importantly, in real life the attractiveness of self-chosen games would be much higher, causing them to possibly be more disruptive.

5 Conclusion

In this study, no detrimental effect of parallel use of an easy gaming app on the learning performance of university students reading a text was observed. Participants' reading time and quiz performance appeared unimpaired and were also not significantly affected by the presence or absence of push notifications. However, for none of these comparisons could statistical equality of the groups be established. This confirms that it is worthwhile to investigate further the effects of smartphone games on study-related tasks. This study is a good starting point for this endeavour and highlights a wide range of relevant directions of experimental research.

6 Directions for future research

The current sample was a homogeneous sample of high achieving students with good impulse control. It will be relevant to investigate the effects in samples with different characteristics, such as persons with a broader range of academic ability, higher impulsivity and less ability to delay gratification. In addition, the attractiveness of the game and the learning task should be varied systematically in order to investigate variations of the learners' (relative) motivation. The gaming task's attractiveness could be increased if participants were highly invested in the game, for example when having spent much time to achieve high scores. To explore whether the present experiment was influenced by the use of compensation strategies, the length of the text should be varied for the costs of compensation to become potentially more apparent.

Acknowledgements The authors wish to thank Kristina Sophie Buch and Shari Jana Dudda for their help in collecting the data and Patrick Hay for his help with data evaluation.

Funding Open Access funding enabled and organized by Projekt DEAL.

Declarations

Conflict of interest The authors do not have any conflicts of interest to report.

Ethical approval The study was conducted in accordance with the Declaration of Helsinki and approved by the internal review board of XXX University (approval ID: XXX). All participants received full information about the study and provided informed consent. Being university students, they received course credits for their participation.

Open Access This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by/4.0/>.

References

- Adler, L. A., Faraone, S. V., Spencer, T. J., Michelson, D., Reimherr, F. W., Glatt, S. J., Marchant, B. K., & Biederman, J. (2008). The reliability and validity of self- and investigator ratings of ADHD in adults. *Journal of Attention Disorders*, 11(6), 711–719. <https://doi.org/10.1177/1087054707308503>
- Arkes, H. R., & Blumer, C. (1985). The psychology of sunk cost. *Organizational Behavior and Human Decision Processes*, 35(1), 124–140. [https://doi.org/10.1016/0749-5978\(85\)90049-4](https://doi.org/10.1016/0749-5978(85)90049-4)
- Borragán, G., Slama, H., Bartolomei, M., & Peigneux, P. (2017). Cognitive fatigue: A time-based resource-sharing account. *Cortex*, 89, 71–84. <https://doi.org/10.1016/j.cortex.2017.01.023>

- Cades, D. M., Werner, N. E., Boehm-Davis, D. A., & Arshad, Z. (2010). What makes real-world interruptions disruptive? Evidence from an office setting. *Proceedings of the Human Factors and Ergonomics Society*, 1(2008), 448–452. <https://doi.org/10.1518/107118110X12829369202475>
- Christiansen, H., Kis, B., Hirsch, O., Philipsen, A., Henneck, M., Panczuk, A., Pietrowsky, R., Hebebrand, J., & Schimmelmann, B. G. (2011). German validation of the Conners Adult ADHD Rating Scales–self-report (CAARS-S) I: Factor structure and normative data. *European Psychiatry*, 26(2), 100–107.
- Conard, M. A., & Marsh, R. F. (2014). Interest level improves learning but does not moderate the effects of interruptions: An experiment using simultaneous multitasking. *Learning and Individual Differences*, 30, 112–117. <https://doi.org/10.1016/j.lindif.2013.11.004>
- Conners, C. K., Erhardt, D., Epstein, J. N., Parker, J. D. A., & Sitarenios, G. (1999). Self-ratings of ADHD symptoms in adults I: Factor structure and normative data. *Journal of Attention Disorders*, 3(3), 141–151. <https://doi.org/10.1177/108705479900300304>
- Couffe, C., & Michael, G. A. (2017). Failures due to interruptions or distractions: A review and a new framework. *American Journal of Psychology*, 130(2), 163–181. <https://doi.org/10.5406/amerjpsyc.130.2.0163>
- Crawford, E. R., LePine, J. A., & Rich, B. L. (2010). Linking job demands and resources to employee engagement and burnout: A theoretical extension and meta-analytic test. *Journal of Applied Psychology*, 95(5), 834–848.
- Czerwinski, M., Cutrell, E., & Horvitz, E. (2000). Instant messaging and interruption: Influence of task type on performance. *Proceedings of OZCHI, 2000*, 356–361. [https://doi.org/10.1016/S1361-3723\(02\)01112-0](https://doi.org/10.1016/S1361-3723(02)01112-0)
- Destatis. (2018). Ausstattung privater Haushalte mit Informations- und Kommunikationstechnik—Deutschland. Retrieved November 19, 2019, from https://www.destatis.de/DE/Themen/Gesellschaft-Umwelt/Einkommen-Konsum-Lebensbedingungen/Ausstattung-Gebrauchsgueter/_inhalt.html
- 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>
- Erhardt, D., Epstein, J. N., Conners, C. K., Parker, J. D. A., & Sitarenios, G. (1999). Self-ratings of ADHD symptoms in adults II: Reliability, validity, and diagnostic sensitivity. *Journal of Attention Disorders*, 3(3), 153–158. <https://doi.org/10.1177/108705479900300304>
- Faber, L. G., Maurits, N. M., & Lorist, M. M. (2012). Mental fatigue affects visual selective attention. *PLoS ONE*, 7(10), 1–10. <https://doi.org/10.1371/journal.pone.0048073>
- Fullwood, C., Quinn, S., Kaye, L. K., & Redding, C. (2017). My virtual friend: A qualitative analysis of the attitudes and experiences of Smartphone users: Implications for Smartphone attachment. *Computers in Human Behavior*, 75, 347–355. <https://doi.org/10.1016/j.chb.2017.05.029>
- Gillie, T., & Broadbent, D. (1989). What makes interruptions disruptive? A study of length, similarity, and complexity. *Psychological Research Psychologische Forschung*, 50(4), 243–250. <https://doi.org/10.1007/BF00309260>
- Graben, K., Doering, B. K., Jeromin, F., & Barke, A. (2020). Problematic mobile phone use: Validity and reliability of the Problematic Use of Mobile Phone (PUMP) Scale in a German sample. *Addictive Behaviors Reports*, 12(July), 100297. <https://doi.org/10.1016/j.abrep.2020.100297>
- Gupta, N., & Irwin, J. D. (2016). In-class distractions: The role of Facebook and the primary learning task. *Computers in Human Behavior*, 55, 1165–1178. <https://doi.org/10.1016/j.chb.2014.10.022>
- Gutt, C., Press, W., Bohrmann, G., & Greinert, J. (2001). Brennendes Eis. *Physikalische Blätter*, 57(9), 49–54.
- Hess, S. M., & Detweiler, M. C. (1994). *Training to reduce the disruptive effects of interruptions*. 38(18), 1173–1177. <https://doi.org/10.1177/154193129403801806>
- Ipsos. (2016). Jeder vierte Deutsche spielt täglich Mobile Games. Retrieved from <https://www.ipsos.com/de-de/jeder-vierte-deutsche-spielt-taeglich-mobile-games>
- Junco, R., & Cotten, S. R. (2012). The relationship between multitasking and academic performance. *Computers and Education*, 59(2), 505–514. <https://doi.org/10.1016/j.compedu.2011.12.023>
- Karpinski, A. C., Kirschner, P. A., Ozer, I., Mellott, J. A., & Ochwo, P. (2013). An exploration of social networking site use, multitasking, and academic performance among United States and European university students. *Computers in Human Behavior*, 29(3), 1182–1192. <https://doi.org/10.1016/j.chb.2012.10.011>
- Katidioti, I., Borst, J. P., Van Vugt, M. K., & Taatgen, N. A. (2016). Interrupt me: External interruptions are less disruptive than self-interruptions. *Computers in Human Behavior*, 63, 906–915. <https://doi.org/10.1016/j.chb.2016.06.037>

- King, D. L., & Delfabbro, P. H. (2014). The cognitive psychology of Internet gaming disorder. *Clinical Psychology Review*, 34(4), 298–308. <https://doi.org/10.1016/j.cpr.2014.03.006>
- Kreifeldt, J. G., & McCarthy, M. E. (1981). *Interruption as a test of the user-computer interface*. 655–667. <https://doi.org/10.1360/zd-2013-43-6-1064>
- Kuznekoff, J. H., & Titsworth, S. (2013). The impact of mobile phone usage on student learning. *Communication Education*, 62(3), 233–252. <https://doi.org/10.1080/03634523.2013.767917>
- Lakens, D. (2017). Equivalence tests: A practical primer for t tests, correlations, and meta-analyses. *Social Psychological and Personality Science*, 8(4), 355–362. <https://doi.org/10.1177/1948550617697177>
- McFarlane, D. (1999). Coordinating the interruption of people in human–computer interaction. In *30th International Conference on Human–Computer Interaction*, (Ntsb 1988), 295–303. <https://doi.org/10.1207/S15327051HCI1701>
- Merlo, L. J., Stone, A. M., & Bibbey, A. (2013). Measuring problematic mobile phone use: Development and preliminary psychometric properties of the PUMP Scale. *Journal of Addiction*, 2013(May), 1–7. <https://doi.org/10.1155/2013/912807>
- Oulasvirta, A., & Saariluoma, P. (2004). Long-term working memory and interrupting messages in human–computer interaction. *Behaviour and Information Technology*, 23(1), 53–64. <https://doi.org/10.1080/01449290310001644859>
- Rosen, L. D., Carrier, L. M., & Cheever, N. A. (2013). Facebook and texting made me do it: Media-induced task-switching while studying. *Computers in Human Behavior*, 29(3), 948–958. <https://doi.org/10.1016/j.chb.2012.12.001>
- Trafton, J. G., & Monk, C. A. (2007). Task interruptions. *Reviews of Human Factors and Ergonomics*, 3(1), 111–126. <https://doi.org/10.1518/155723408X299852>
- VuMa. (2019). Studierende in Deutschland nach Marke des eigenen Handys oder Smartphones im Jahr 2019. Retrieved from In Statista website: <https://de.statista.com/statistik/daten/studie/860472/umfrage/umfrage-unter-studenten-zu-den-beliebtesten-marken-bei-smartphones/>
- Walsh, S. P., White, K. M., & Young, R. M. (2010). Needing to connect: The effect of self and others on young people’s involvement with their mobile phones. *Australian Journal of Psychology*, 62(4), 194–203. <https://doi.org/10.1080/00049530903567229>
- Widmer, P. S., Semmer, N. K., Kälén, W., Jacobshagen, N., & Meier, L. L. (2012). The ambivalence of challenge stressors: Time pressure associated with both negative and positive well-being. *Journal of Vocational Behavior*, 80(2), 422–433. <https://doi.org/10.1016/j.jvb.2011.09.006>
- Wood, E., Zivcakova, L., Gentile, P., Archer, K., De Pasquale, D., & Nosko, A. (2012). Examining the impact of off-task multi-tasking with technology on real-time classroom learning. *Computers and Education*, 58(1), 365–374. <https://doi.org/10.1016/j.compedu.2011.08.029>

Publisher’s Note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.