Computers & Education 141 (2019) 103611

Contents lists available at ScienceDirect

Computers & Education

journal homepage: www.elsevier.com/locate/compedu

Understanding smartphone usage in college classrooms: A long-term measurement study

Inyeop Kim[a], Rihun Kim[a], Heepyung Kim[a], Duyeon Kim[a], Kyungsik Han[c], Paul H. Lee[b], Gloria Mark[d], Uichin Lee[a,*]

[a] *Korea Advanced Institute of Science and Technology, Republic of Korea*
[b] *Hong Kong Polytechnic University, Hong Kong*
[c] *Ajou University, South Korea*
[d] *University of California, Irvine, USA*

ARTICLE INFO

Keywords:
In-class smartphone use
Multitasking behaviors
Objective measurements
Smartphone distraction
Academic performance

ABSTRACT

Smartphone usage is widespread in college classrooms, but there is a lack of measurement studies. We conducted a 14-week measurement study in the wild with 84 first-year college students in Korea. We developed a data collection and processing tool for usage logging, mobility tracking, class evaluation, and class attendance detection. Using this dataset, we quantify students' smartphone usage patterns in the classrooms, ranging from simple use duration and frequency to temporal rhythms and interaction patterns. Furthermore, we identify the key predictors of students' in-class smartphone use and their semester grades. Our results reveal that students use their phones for more than 25% of effective class duration, and phone distractions occur every 3–4 min for over a minute in duration. The key predictors of in-class smartphone use are daily usage habits and class characteristics, and in-class phone usage is negatively correlated with student grades.

1. Introduction

Smartphone use among college students is becoming increasingly prevalent even in classrooms. According to the cognitive theory of multimedia learning, off-task multitasking while learning (e.g., texting and social media use) could interfere with a learner's information processing, thereby lowering learning performance (Bowman, Levine, Waite, & Gendron, 2010; Kuznekoff & Titsworth, 2013). Experimental results confirm the negative influence of off-task multitasking (Waite, Lindberg, Ernst, Bowman, & Levine, 2018). Recent survey studies further showed that overall academic performance is negatively related to social media multitasking (Junco & Cotten, 2012) and smartphone addiction (Lepp, Barkley, & Karpinski, 2015; Samaha & Hawi, 2016).

Despite growing concerns, there is a lack of knowledge regarding how students actually multitask with their phones in classrooms and what the predictors of in-class smartphone use are. Existing studies are mostly based on controlled experiments and survey methods (Junco & Cotten, 2012; McCoy, 2016), with which it is difficult to acquire accurate and naturalistic usage behaviors due to recall errors and social desirability bias of under-reporting undesirable behaviors. Alternatively, several recent studies adopted automated sensing and logging for objective data analysis. The StudentLife project (R. Wang et al., 2014; R. Wang, Harari, Hao, Zhou, & Campbell, 2015) used mobile sensing (e.g., activity, location) to investigate students' mental health and academic performance, and the EDUM project proposed to analyze campus Wi-Fi traffic patterns to infer students' in-class behaviors (e.g., arrival/departure and

* Corresponding author. Korea Advanced Institute of Science and Technology, 291 Daehak-ro, Yuseong-gu, Daejeon, 34141, Republic of Korea.
E-mail address: uclee@kaist.ac.kr (U. Lee).

https://doi.org/10.1016/j.compedu.2019.103611
Received 11 March 2019; Received in revised form 17 June 2019; Accepted 20 June 2019
Available online 26 June 2019
0360-1315/

network activities) (Zhou et al., 2016). Several studies examined short-term measurement data (typically less than a month) to understand the multitasking behaviors of college students (Mark, Wang, & Niiya, 2014; Y.; Wang, Niiya, Mark, Reich, & Warschauer, 2015). These approaches formed the foundations for using advanced sensing and measurements for understanding students' multitasking behaviors in classrooms. However, none of the studies examined the fine-grained in-class usage patterns, the predictors of in-class phone use, and the impacts of phone use on grades. These will provide valuable insights into student modeling, educational decision making, and pedagogical support (Romero & Ventura, 2010).

We conduct a long-term measurement study with 84 students in the Fall semester of 2017 in Korea. We collected 196,704 h of smartphone usage data for 14 weeks as well as class-specific self-reported data of 445 classes over the semester. We used a simple algorithm to check class attendance and student arrival/departure. Since a majority of lectures last 90 min, we limited our analysis on the 13,608 h of in-class usage data from 353 courses in classes of duration 90 min. Specifically, this dataset allows us to answer the following research questions.

- RQ1) What are the measured (and self-reported) usage patterns of (in-class vs. overall) smartphones among students?
- RQ2) What kinds of apps do students frequently use in class and how do they interact with them?
- RQ3) What are the temporal patterns of smartphone use during 90-min classes?
- RQ4) What are the predictors of in-class smartphone use? (e.g., overall usage characteristics, course difficulty, class size, and lecture organization)
- RQ5) How are smartphone usage predictors related to semester grades? (e.g., overall or in-class usage characteristics).

Our results showed that our student participants use their phones for more than 20 min per class on average (M = 21.06 min, SD = 9.22). Given that the actual duration of 90-min lectures after excluding typical 15-min breaks is 75 min, their usage is over 25% of the actual class duration. The average usage frequency per class was 12.28 (SD = 5.27). There was considerable underestimation regarding how frequency students used their phones in classrooms (self-report: 8.25 vs. measurement: 12.28). No significant differences in frequently used apps were found between in-class and overall usage datasets, and yet in both cases, instant messaging, social media, and web browsers were dominantly used as opposed to other apps. A majority of user interactions in classrooms were scrolling and typing, and notifications were by far largely from instant messaging. Temporal usage analyses showed that smartphone usage rapidly ramped up as soon as the class started, and phone distractions happened in every 3–4 min for over a minute. Our multilevel regression results showed that in-class usage duration was related to daily usage and class size, whereas in-class usage frequency was related to daily usage frequency and lecture organization. Furthermore, smartphone use was indeed closely related to the grades, and the major predictors include daily usage duration and various in-class usage of apps (i.e., instant messaging and web surfing). Finally, we discuss the prevalence of in-class smartphone usage, its influence on attention duration and academic performance, the methodological contributions, and intervention design.

There is a growing trend of adopting digital technologies in classrooms (e.g., smart classrooms, intelligent tutoring), and thus, it is increasingly important to have an in-depth understanding of technology use. The major contribution of this work is to collect and analyze a long-term real-usage dataset in the wild. Our work significantly expands the current scholarship on digital technology use behaviors in educational contexts (Junco & Cotten, 2012; McCoy, 2016; Y.; Wang et al., 2015) by identifying in-class phone usage patterns, the predictors of in-class phone use, and the impacts of phone use on grades. Our results provide novel insights into student modeling (e.g., technology use, distraction patterns), educational decision making (e.g., in-class technology use policy and computer-supported learning), and pedagogical support (e.g., designing learning strategies and tools) (Romero & Ventura, 2010), which can also be applicable to intelligent tutoring or computer-supported learning environments. Furthermore, unlike existing measurement studies (R. Wang et al., 2014; Zhou et al., 2016), our measurement framework brings new opportunities for data analysis to the fields of educational data mining and learning analytics (Romero & Ventura, 2010).

2. Background and related work

2.1. Information processing and multitasking

As the cognitive theory of multimedia learning states, information processing for learning involves multiple channels such as auditory/verbal and visual/pictorial inputs (Mayer & Moreno, 2003). Each channel has limited processing capacity as the multiple resource theory states (Wickens, Hollands, Banbury, & Parasuraman, 2015). Learning typically requires considerable cognitive processing over these channels; e.g., information from each channel is selected and organized in the working memory, which is then integrated into existing knowledge in the long-term memory. Multitasking while learning (e.g., texting and social media use) may interfere with a learner's information processing, negatively influencing unit learning tasks such as reading, note-taking, and recalling (Bowman et al., 2010; Kuznekoff & Titsworth, 2013). Furthermore, researchers found that heavy multitaskers had difficulty in filtering out irrelevant information and were slower in switching tasks than light multitaskers (Ophir, Nass, & Wagner, 2009). More detailed relationships between digital media and cognitive functions can be found in a recent review (Wilmer, Sherman, & Chein, 2017).

Existing human-computer interaction studies investigated multitasking behaviors in various contexts. Individuals organize their work spheres into related tasks and frequently switch between different spheres (González & Mark, 2004), and a typical work fragment lasts for only approximately 11 min (Mark, González, & Harris, 2005). Multitasking is mostly due to external interruptions such as phone calls and emails; however, 40% of task switching is self-initiated (Czerwinski, Horvitz, & Wilhite, 2004). Prior work showed that self-initiated multitasking happens owing to adjustments, habitual routines, and triggers/stimuli (Jin & Dabbish, 2009), and this helps to maintain the flow state of work (e.g., seeking for or staying away from challenging tasks) (Adler & Benbunan-Fich, 2013). A user's attentional states are related to various contextual factors such as online activities, the time of day, and day of the week (Mark et al., 2014). Several studies analyzed users' multitasking management practices; e.g., local vs. remote social interactions (Ames, 2013), multiple social interaction channels (Wohn & Birnholtz, 2015), and work-life balance with boundary management (Lim, Arawjo, Xie, Khojasteh, & Fussell, 2017).

2.2. Digital distractions and off-task multitasking

Potential sources of distraction in traditional classrooms include various internal and external distractors (Tesch, Coelho, & Drozdenko, 2011). External distractors include difficulties in understanding learning materials and instructors' teaching in class, chatting noise, and technology use of other students (e.g., phone ringing, laptop noise). There are also well-known internal distractors such as illness, drowsiness, and personal technology use (e.g., phone ringing, gaming, music, texting, email checking). When students are distracted, they may shift their attention to mobile phones as a coping strategy (e.g., avoiding boring lectures by checking Facebook updates).

Off-task multitasking, such as texting in class is related to usage habits and media gratifications (e.g., pleasure, escape, affection, inclusion, and relaxation) (Wei & Wang, 2010). For example, college students habitually use text messaging to chat with their friends to cultivate their interpersonal relationships. Prior work also showed that off-task multitasking during college classes is mostly attributed to texting, Facebook, email, and web search, and the texting usage ratio is far greater than that of the others (texting: 69% vs. Facebook: 28%) (Junco & Cotten, 2012). According to recent survey results in US universities (McCoy, 2016), the major motives for in-class phone use include boredom, fun, social connection, urgency, and class purpose. The off-task usage frequency in class was 11.43 times per class on average. Undergraduate students use their phones more frequently than graduate students for non-class activities.

Digital distractions can also happen in computer-based learning environments (e.g., classrooms with intelligent tutors) or e-learning environments (e.g., massive open online classes). When students engage in intelligent tutors, researchers found that students experience frequent cognitive-affective states such as boredom and confusion, which may lead to off-task behaviors (Baker, D'Mello, Rodrigo, & Graesser, 2010). Besides the existing distraction sources (Tesch et al., 2011), learning motives and attitudes (e.g., topical liking, self-motivation) are closely related to off-task behaviors (Baker, 2007).

2.3. Digital distractions and academic performance

Off-task multitasking negatively influences student engagement (Junco & Cotten, 2012) and learning performance (Gingerich & Lineweaver, 2014; Wood et al., 2012). When texting is considered, the results of controlled experiments showed that texting had a negative influence on test scores (Gingerich & Lineweaver, 2014) and the quality of note taking (Waite et al., 2018). Likewise, off-task multitasking, such as Facebook and text messaging during classroom lectures negatively influences learning performance (Wood et al., 2012). In-class laptop use is a significant distractor to both users and nearby students (Fried, 2008; Sana, Weston, & Cepeda, 2013), and laptop use was negatively associated with learning performance (Gaudreau, Miranda, & Gareau, 2014; Ravizza, Uitvlugt, & Fenn, 2017; Wood et al., 2012).

Prior survey studies of off-task digital media use mostly reported negative effects as well, although several early studies reported that there were no significant differences (e.g., Facebook use vs. grade point average (GPA)) (Josh Pasek, Eian More & Eszter Hargittai, 2009; Kolek & Saunders, 2008). The amount of cellphone use and its dependence are negatively correlated with the academic performance (Lepp et al., 2015; Samaha & Hawi, 2016). In addition, off-task multitasking while studying (e.g., social media and texting) is positively correlated with smartphone dependence (David, Kim, Brickman, Ran, & Curtis, 2015). Junco and Cotton (Junco & Cotten, 2012) conducted a large-scale survey study (n = 1839) to investigate the relationship between off-task multitasking and GPA. They found that Facebook and texting were negatively associated with GPA, whereas emailing and talking on the phone were not. Furthermore, researchers found that social media usage differs across class ranks (Junco, 2015): students with lower class ranks (freshmen, sophomores, and juniors) spent significantly more time on Facebook (and more time on multitasking with Facebook) than senior students, and negative impacts of off-task multitasking on GPA were found among students with lower class ranks. A recent study emphasized the importance of the direction through a path analysis, showing that students' GPAs influence how much time they spend on social media, though not vice versa (Michikyan, Subrahmanyam, & Dennis, 2015).

There is still a lack of real-world measurement reports because existing studies are mostly based on controlled experiments and survey studies (Junco & Cotten, 2012; McCoy, 2016). A mixed study of observing and surveying 263 middle school, high school, and university students for 15 min in their homes showed that in less than 6 min, students were distracted by their phones owing to social media and texting (Rosen, Carrier, & Cheever, 2013). Earlier measurement studies focused on understanding generic smartphone usage patterns (Böhmer, Hecht, Schöning, Krüger, & Bauer, 2011; Falaki et al., 2010; Tossell, Kortum, Shepard, Rahmati, & Zhong, 2014). The StudentLife project (R. Wang et al., 2014, 2015) leveraged smartphone sensing (i.e., activity, location, conversation, sleep) and experience sampling (i.e., stress, depression) to collect a real-world dataset from 48 college students for 10 weeks. The project uncovered that students' behavioral characteristics (e.g., conversation, mobility) and their trends (e.g., up/down directions, change in directions) are related to mental health and academic performance. Researchers demonstrated that monitoring campus Wi-Fi network traffic helps to understand students' behaviors (e.g., attendance, arrivals/departures, phone use) (Zhou et al., 2016). Several studies examined short-term measurement data to understand specific aspects of multitasking behaviors (e.g., social media

Table 1
Smartphone usage and sensor data.

Data type	Description
Activity	Activity class in every 15s (i.e., Still, Walk, Run, Bike, Vehicle)
Application	Installed/running apps
Data Traffic	Number of bytes transmitted or received (every 10s after Screen On)
GPS	GPS location in every 5s if activity is not still
Wi-Fi Fingerprint	Wi-Fi scanning and currently associated AP if activity is still
Notification	Notification source, title, alarm type (LED/Vibration/Sound), notification setting
Ringer Mode	Current ringer mode (Silence/Vibration/Sound
Screen On/Off	Screen On/Off event
Touch	Screen touch (short/long/scroll)
Keyboard	Key press event

checking patterns (Y. Wang et al., 2015) and the relationship between stress and multitasking (Mark et al., 2014)) and to analyze the impact of social media use on grades (Y. Wang & Mark, 2018). None of the existing studies on phone usage measurements considered understanding multitasking behaviors in classrooms. Our work extends these studies and aims to study detailed multitasking patterns in classrooms and their relationship on academic performance.

3. Data collection and preprocessing

3.1. Data collection method

We developed a mobile app for smartphone usage and sensor data collection, and a web-based portal for survey data collection. The details are shown in Tables 1 and 2. Smartphone usage and sensor data collection were implemented based on Android's accessibility service, which helped us to collect the interaction and sensor data in the background. As shown in Table 1, the data types include user activity, app use (installed and running), data traffic (send/receive), GPS, Wi-Fi fingerprints, notifications, ringer mode, screen on/off, and touch/keyboard events. Our app temporarily stores all the collected data as an SQL file in the phone's local directory. We used DropSync to automatically synchronize stored data to the researcher's Dropbox account. The data transfer was scheduled in every 6 h, and actual data transfer happens only when the phone was connected to the Wi-Fi network to minimize cellular data transfers.

We also collected self-reported data related to class evaluation. Every week each student was asked to rate three items on a 5-point Likert scale (i.e., difficulty, interest, and workload) and one item on a binary scale (i.e., whether there was a quiz/exam) for each course. Furthermore, each student evaluated the course every month based on the university's three-item course evaluation questionnaire on a 5-point Likert scale: organization of a class, instructor's delivery effectiveness, and the class's helpfulness for learning. We also collected a battery of psychological surveys such as personality, stress, and self-esteem, which were not considered in the current analysis. Given that commercial survey software does not allow customized survey generations per user, we designed our own web-based portal for survey data collection and user monitoring (for checking data collection status).

After the data collection period, we collected data from several other survey instruments. First, we collected self-reported in-class on- and off-task smartphone usage (frequency per 90-min class, and duration of each use) and usage purposes (app names and usage reasons). Second, we asked users to complete the smartphone addiction scale (SAS) (Kwon et al., 2013). Third, we asked the students to self-report their course grades and the overall grade point average (GPA) for the semester by referring to the university systems.

3.2. Procedure and participants

We collected data during the Fall semester of 2017 at a large technical university in Korea. From September 3 to 20, we invited

Table 2
Class related data, self-reported usage, and grades.

Data Type	Description	Items	Period
Course Info.	Course information: time, location, enrolled students, credits		Pre
Weekly Report	Difficulty/Interest/Workload (5-point Likert scale), Quiz/Exam (binary scale)	4 items per class	Weekly
Monthly Report	Organization, delivery effectiveness, and helpfulness for learning (5-point Likert scale)	3 items per class	Monthly
Self-reported Usage	On-task usage frequency per 90 min class(+ duration per use) + On-task usage purpose (multiple selection), Off-task usage frequency per 90 min lecture (+ duration per use) + Usage purpose (multiple selection)	6 items	Post
Grades	Letter grade of each course and overall semester GPA		Post

students by announcing our study on the campus community site. Given that prior studies showed that off-task multitasking (e.g., social media) differs across class ranks (Junco, 2015), we controlled the class ranks by focusing on freshmen. We recruited 92 participants who were first-year students and Android users. We invited the participants to the orientation session and explained the study goals, procedures, and requirements, and answered any questions that they had. Eighty-one participants started their data collection on September 10th (for 14 weeks), and the remaining 11 participants started on September 22nd (for 12.3 weeks). Eight participants dropped out of the data collection, and a total of 84 participants completed the study. The reasons for dropouts include data collection burdens (n = 6), smartphone change (to the iPhone) (n = 1), and leave of absence (n = 1). Among the 84 participants, 56 (67%) were males, and the average age of participants was 19.61 (SD = 0.6). The skewed gender distribution was due to the fact that the study was conducted in a large technical university where about 70% of the students are male. We used our web-based portal to track the data collection status by checking the last data sync time. If data were missing for longer than one day, we worked with the participants to resume data collection. Our study was reviewed and approved by the university's Institutional Review Board (IRB) and was conducted with all participants' written consent to collect and use the information. Participants who completed all study requirements were compensated with 100 USD (50 USD for dropouts).

Our participants' class enrollment status was given as follows: 4 courses (n = 11), 5 courses (n = 34), 6 courses (n = 34), 7 courses (n = 13), and 8 courses (n = 2). Their credits showed the following distribution: 10–12 (n = 8), 13–15 (n = 32), 16–18 (n = 36), 19–21 (n = 3), and 22–24 (n = 2). The number of unique courses was 445, and their class duration distribution was given as follows: 90 min (353 courses), 120 min (30 courses), 180 min (60 courses) and 210 min (2 courses). Given that the usage duration and frequency per class are dependent on class duration, we focus on the 90-min classes, which account for 80% of all the classes. Note that the 90-min classes usually last 75 min due to 10–15 min of breaks between classes. By referring to the registrar's database, we found that the average enrollment size of the 353 90-min classes was 83.55 (SD = 43.12). Overall, we collected 196,704 h of smartphone usage from 84 students over 14 weeks, and 13,608 h of in-class usage data from 90-min classes were analyzed to understand the usage patterns.

3.3. Data preprocessing

Given that our study focuses on the use of smartphones during lectures, we consider the participants' attendance in the lecture through the use of GPS and Wi-Fi fingerprint information as well as the detection of the classroom arrival/departure based on activity data.

To detect class attendance, we used Wi-Fi fingerprints (i.e., MAC address, received signal strength) and GPS coordinates (latitude and longitude). As in prior studies (Bahl & Padmanabhan, 2000; Zhou et al., 2016), the collected Wi-Fi fingerprints and GPS coordinates are compared with each classroom's fingerprints. One of the authors visited every classroom on campus and built this fingerprint database. Collecting a fingerprint database is a laborious task. We found that campus Wi-Fi hotspots have spatial locality in their MAC addresses. This property was exploited to approximate the coverage of indoor localization. If Wi-Fi fingerprints are missing, fingerprint matching fails; we found that users sometimes turned off their Wi-Fi interfaces for battery saving, although we instructed them not to do so. In such cases, we alternatively checked whether there were any GPS logs for classroom detection. If none of these data were available, we could not check attendance, and thus, we excluded such log data in our analysis.

In the case of GPS coordinates (latitude and longitude), the position, where the strength of the GPS signal gradually weakens and is eventually disconnected, is regarded as a point to distinguish between outdoors and indoors (Ashbrook & Starner, 2003). The time when the GPS data are missing is considered as the time when the student entered a building and the time when the GPS data start being collected again is considered as the time when the student exited the building; the difference between these two times is considered as the time inside the building. We select the building that was closest to the last value of the GPS data. However, we found that in some cases GPS readings were sporadically collected even within the buildings, and users sometimes turned off GPS sensors in the building. To address this issue, we manually created the circles that signify each building's boundary, and crossing the boundary denotes arrival and departure. The size of a circle is determined such that it can enclose a building's circumference. We considered that a user crosses the boundary line, if the distance between the center coordinate of the circle and the user's current coordinate is smaller than the radius of the circle. A circular shape was preferred because even if a user enters the building from any direction, boundary crossing is always based on the length from the origin.

We verified the reliability of our attendance checking method as follows. After the semester ended, we received a total of 628 attendance records of 13 participants from the internal university learning system and compared them with the detected attendance results from our application—only limited attendance information is available because instructors or TAs must manually enter attendance information. We excluded 89 cases without proper sensor data (about 14.17%). Among 539 cases, we found that the detected attendance results and submitted attendance records matched 100%, showing that our method can reliably infer class attendance.

Once the attendance was checked we then detected classroom arrival and departure using activity data. After entering a building, students arrive at the classroom by walking or running and then stay seated, remaining stationary. Likewise, we can safely assume that they depart from the classroom when their activity is changed to walking or running from stationary. Note that we did not use periodic Wi-Fi scanning owing to battery consumption during data collection (D. H. Kim, Kim, Estrin, & Srivastava, 2010), and GPS readings were not used owing to large errors (these are only feasible to detect building-level arrival/departure). In our work, we considered the activity log data from 15 min before and after the lecture time; assuming 15 min breaks between lectures, we used 75 min as the *effective class duration*. Classroom arrival and departure events are basically defined by tracking changes in activity (Cho, Lee, Noh, Park, & Song, 2015). Prior studies on energy efficient activity tracking systems used the time threshold of 90 s to

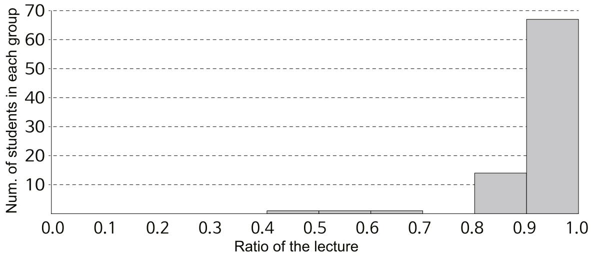

Fig. 1. Histogram of the ratio of the lectures during which students used the phone.

determine whether a user becomes stationary (so that the systems can switch off GPS for battery conservation) (Paek, Kim, & Govindan, 2010; Ben Abdesslem, Phillips, & Henderson, 2009). By referring to the prior studies and manually examining the dataset, we empirically set the time threshold of 90 s.

3.4. Usage analysis model

We call an instance of smartphone use as a *session*. A session starts when the screen is on (i.e., a user clicks the power/home button, or a user reacts to a notification received), and it ends when the screen is off. For a given session, a user can run a series of apps. For example, a user checks notifications in an instant messaging app and then runs Facebook to read news feeds. The duration of a given session is called *session duration.* The time interval between two consecutive sessions is called *inter-session duration.* The number of sessions is simply referred to as *use frequency*, and the sum of all session durations is called *use duration*. These metrics can be calculated in two levels: day- and class-level (e.g., daily usage duration vs. usage duration per class).

4. Results

4.1. RQ1: overall usage statistics

In-class individual smartphone use prevalence: We first checked what the ratio of classes was in which a student uses a phone at least once per 90-min class. Here, we did not consider how much or how frequently students used their phones. When the student always uses the phone in all the classes throughout the semester, the ratio is 1. In contrast, when the student does not use the phone at all, the ratio is 0. As shown in Fig. 1, we found that the majority of students used their phone in more than 90% of their classes (M = 0.93, SD = 0.09). Only 3 students used their phone in less than 80% of their classes. Overall, smartphone use during class is quite prevalent among students.

In-class smartphone usage: We then calculated in-class smartphone usage. We first examined the mean session duration and inter-session duration. The mean session duration tells us the average amount of time spent per phone use, whereas the mean inter-session duration refers to the mean time interval between consecutive sessions during which the student does not use the phone. It is also very interesting to see whether the overall usage patterns over a day differ from those during the classes.

Fig. 2 shows the results. The mean session duration was 117.52s (SD = 61.58), and the inter-session duration was 215.88s (SD = 101.60). These results roughly indicate that our participants were distracted by their phones once every 3–4 min for over a minute. We also calculated the average use frequency per class, and the average use duration per class. Fig. 3 shows the distribution of the students' phone use frequency and duration per class. The average use frequency per class was 12.28 times (SD = 5.27), and the average use duration per class was 21.06 min (SD = 9.34). Assuming effective class duration of 75 min for 90-min classes, students used their phones for more than a quarter of the class duration (28.1%), and they were distracted in every 3–4 min for over a minute. As shown later, it is likely that in-class smartphone usage is off-task-related, because a majority of apps used during classes are social media and instant messaging. We also compare whether the in-class use differs from the overall usage. The Shapiro-Wilk normality test revealed that the values of session and inter-session durations were not normally distributed ($p < 0.05$). Thus, we used

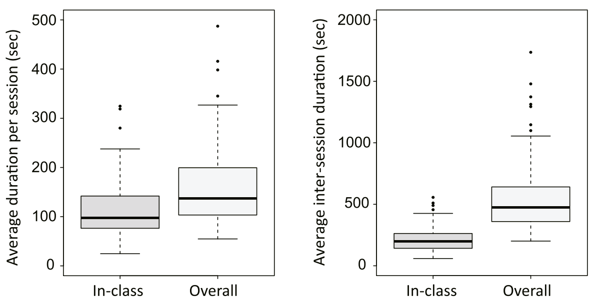

Fig. 2. Average duration per session and inter-session duration.

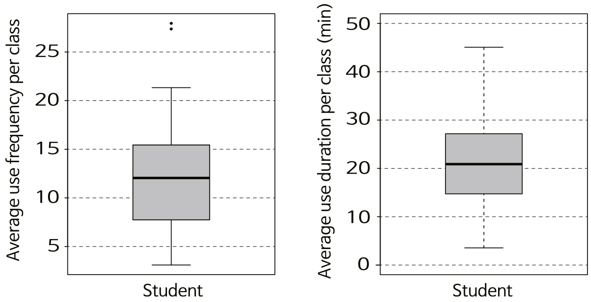

Fig. 3. Average use frequency and duration per class.

the Wilcoxon signed rank test, a non-parametric (or distribution free) alternative to the paired t-test. Both the session and inter-session durations *during classes* are shorter than the *overall* session and inter-session durations with the mean differences of 67.16s and 342.54s, respectively (session duration: Z=-4.11, $p < 0.001$, effect size r=0.32, inter-session duration: Z=-9.66, $p < 0.001$, effect size r= 0.75).

We also examined how the students' phone use duration and frequency during the class changed over the semester (see Fig. 4). The sixth week was removed due to week-long holidays in Korea. We divided the semester into the first and second half of the semester. The first half of semester is from the beginning of data collection to the midterm exam (3–7 week), and the second half of semester is from the end of the midterm exam to the final exam (9–15 week). We analyzed the difference between in-class smartphone use frequency and duration in the first and the second half of the semester, excluding exam periods (week 8 and 16). The duration and frequency of in-class smartphone use were 20.25 min (SD = 17.67) and 12.69 times (SD = 9.71) during the first half of the semester, and 22.06 min (SD = 18.00) and 12.26 times (SD = 9.42) during the second half of the semester, respectively. We compared in-class smartphone use of the first and second half of the semester. Since the Shapiro-Wilk normality test revealed that duration and frequency values were not normally distributed ($p < 0.05$), we used the Mann-Whitney U test, a non-parametric alternative to the independent t-test. There was no significant difference in in-class smartphone use *frequency* in the first and second half of the semester ($p = 0.13$). A significant difference in use duration ($U = 3,419,922$, $Z = -4.01$, $p < 0.001$, effect size $r = 0.05$) was found, indicating that the students tended to use their smartphones longer after midterm exams.

Measured use vs. self-reported use: We asked the participants to self-report how frequently they use their phones for on- and off-task purposes during a typical 90-min class, and how long they use them in each instance of usage (i.e., session) for on- and off-task purposes. Our participants answered that on- and off-task use frequencies were 3.08 (SD = 2.45) and 5.16 times (SD = 3.34),

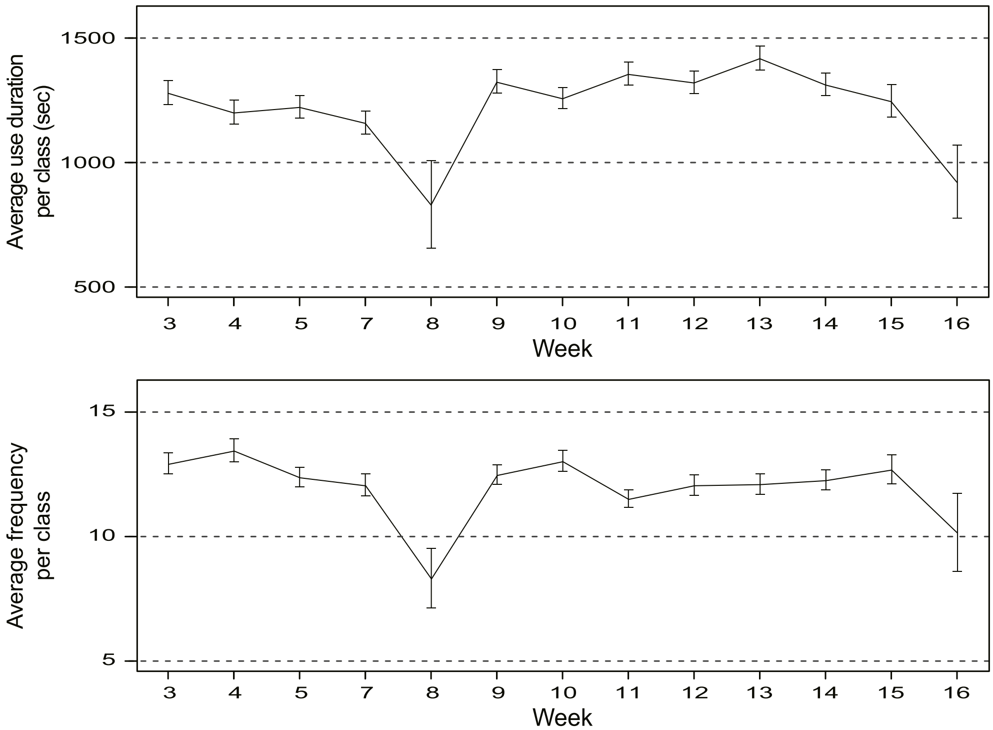

Fig. 4. Use duration and frequency per class over 14 weeks (week 6 was excluded due to holidays).

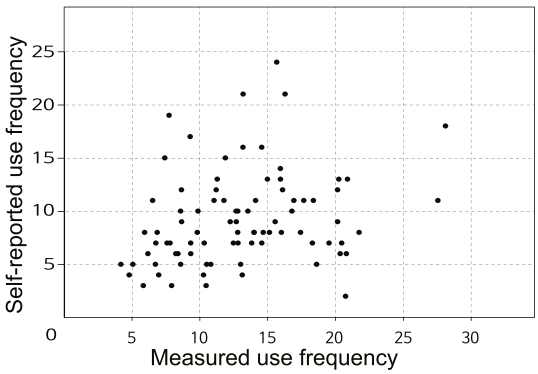

Fig. 5. Use frequency of measured vs. self-reported values.

respectively. The self-reported session durations for on- and off-task use were 293.64s (SD = 226.13) and 395.48s (SD = 386.51), respectively. These values are much greater than the measured session duration (M = 117.52s and SD = 61.58).

We assume that the overall self-reported use frequency is simply the sum of the on- and off-task use frequencies. We found that the overall *self-reported use frequency* per class was 8.08 times (SD = 4.21), whereas the *measured use frequency* per class was 12.29 times (SD = 5.17). Since the distributions of self-reported use frequency and measured use frequency did not follow the normal distribution, we used the Wilcoxon signed-rank test for comparison and calculated the Spearman's rank correlation coefficient. A test shows that there was a significant difference in use frequency ($Z = -7.11$, $p < 0.001$, effect size $r = 0.55$). Fig. 5 shows a scatter plot of the use frequency of self-reported and measured values (Spearman's $\rho = 0.31$, $p < 0.01$). Overall, we find that our participants significantly *underestimated* their use frequency, but *overestimated* session duration per use.

Daily use duration and frequency: We calculated how frequently and how long students use their phone over a day. Our results showed that on average our participants used their phones for 5.78 h (SD = 1.73). The use frequency per day was 164.60 times (SD = 63.60). We also calculated how students use their phone during weekdays and weekends. During weekdays and weekends our participants used their phones for 5.87 h (SD = 1.79) and 5.63 h (SD = 1.67), respectively. The use frequency during weekdays and weekends were 176.12 (SD = 65.89) and 142.64 times (SD = 60.32), respectively. We performed a Shapiro-Wilk normality test to check the normality of the distribution. Use frequency values during weekdays and weekends followed the normal distribution ($p = 0.26$ and $p = 0.22$, respectively), but use duration of weekdays and weekends did not meet the normality assumption ($p < 0.05$). When comparing the use duration and frequency during weekdays and weekends, a Wilcoxon signed-rank test result showed that there was a significant difference in use duration ($Z = -6.02$, $p < 0.001$, effect size $r = 0.46$), and a paired *t*-test result showed that there was a significant difference in use frequency ($t = 3.39$, $p < 0.001$, Cohen's $d = 0.52$). Students used their phones significantly longer and more frequently on weekdays than weekends.

4.2. RQ2: frequent apps and interactions

We found that the student participants used their smartphones for 21.06 min on average (out of 75-min effective class duration), which is more than one-fourth of the class. Based on this surprising result, we examined the smartphone use with respect to the apps, looking into the types of the top apps used and the way the students used them. For the app-based analysis, we analyzed the usage patterns of popular apps among students, and the top 5 popular apps were selected as follows. We first normalized the frequency and duration of each app use for each user. We then summed up the normalized values of all the participants for ranking. Frequently used apps for a long time among many users will be ranked highly according to our metric. However, we found that several students showed overly skewed usage of a few less popular apps (e.g., mobile games and online community apps), which resulted in abnormally high values on these apps. To minimize such a rank bias, we considered the apps that were used above an initial support threshold of 20% based on each user's top 10 apps, which means that at least 20% of participants ($n = 17$) have those apps in their top 10 apps. We manually incremented the support threshold with a gap of 5% and found that 20% was large enough to exclude such apps. The resulting top five apps were given as follows: KakaoTalk (an instant messenger), Facebook, Web browsers (Samsung and Chrome), and Naver (the largest web portal in Korea). These top five apps accounted for 83.9% of the overall usage duration. Fig. 6 shows the boxplots of the top five apps used with respect to the average use duration and frequency; additionally, the university's official learning management system (LMS) app was included for comparison purposes. The LMS app supports various learning activities for the lectures, such as assignment checking, quiz taking, online discussion, Q&A, and lecture material downloading. We found several interesting insights from the results.

Frequently used apps: For both metrics, the instant messaging app (KakaoTalk) ranked the top (average duration: 3.71 min, average frequency: 5.64 times), followed by Facebook (average duration: 2.89 min, average frequency: 2.07 times), and web browsers (average duration: 2.29 min, average frequency: 1.94 times). Since the use duration and frequency of top 5 apps did not satisfy the normality assumptions, we used the Kruskal-Wallis test for group comparison, a non-parametric (or distribution free) alternative to the one-way ANOVA. Post-hoc comparisons were performed using a Bonferroni-Dunn test. The use durations and frequencies of KakaoTalk and Facebook were significantly greater than those of other apps (duration: $F(5, 498) = 49.63$, $\chi^2(x) = 172.02$, $p < 0.001$,

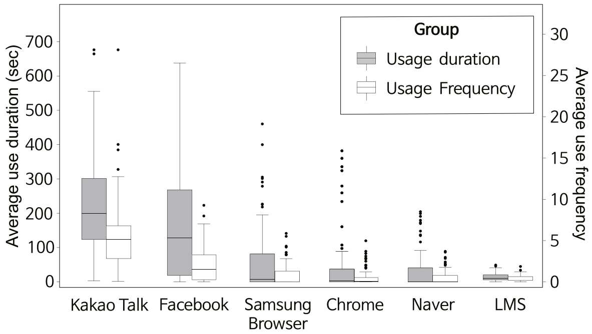

Fig. 6. Use duration and freq. of top 5 apps and LMS app.

frequency: F(5, 498) = 77.10, $\chi^2(x) = 197.89$, $p < 0.001$). In particular, the frequencies of KakaoTalk and Facebook use were much higher than other apps. This indicates that, during the class, it is likely that most students were multitasking by texting with their friends and checking news feeds from social media.

App multitasking: Our student participants ran an average of 2.33 apps per session (SD = 0.69) during class. Note that we showed earlier that the mean session duration was 117.52s (SD = 61.58). This indicates that the students exhibited shorter use of multiple apps, involving brief bursts of interaction with apps (known as micro-usage behaviors (Ferreira, Goncalves, Kostakos, Barkhuus, & Dey, 2014)). It is likely that one app use entails other apps in a given session rather than a single app dominating the session, which is supported by the result that the average number of apps executed per class for each student was 17.15 (SD = 9.33). Frequent app pairs include KakaoTalk/Facebook and KakaoTalk/Browser.

Ringer mode and notifications: We also examined how students set their ringer mode (i.e., sound, vibration, and silent modes) to understand possible interruptions due to notifications during the classes. In Fig. 7, we presented the histogram of ringer mode ratios over all classes per student. The ratio means the fraction of classes in which a student's phone was set to a given mode. As expected, the sound mode was least preferred among all the students. A majority of our participants preferred either silent or vibration mode during the classes; however, there are still quite a few students who use both silent and vibration modes quite often. Given that ringer modes are closely related to notification deliveries as a possible source of external interruptions, we examined how many notifications a user receives for a given application. According to a prior study (U. Lee et al., 2014), a majority of notifications originate from KakaoTalk, Facebook, and the text messaging app. When considering these apps, we found that the mean number of notifications per class from these apps were 13.80 (SD = 16.18), 0.13 (SD = 0.16), and 0.19 (SD = 0.17), respectively. Consistent with prior results (U. Lee et al., 2014), KakaoTalk was the dominant source of notifications even in classrooms. Given that KakaoTalk was one of the major apps used in the classrooms, it is likely that notifications may serve as multitasking triggers.

4.3. RQ3: temporal rhythm of app use in 90-min classes

How does the smartphone use of a student change during a class? As shown earlier, most smartphone use is likely related to off-tasks such as instant messaging and social media use. Digital distractions will considerably influence students' attention levels. Prior studies mostly claim that students' attention levels start declining shortly after a class starts (Bradbury, 2016) or after the peak of their attention levels around 10–15 min (Stuart & Rutherford, 1978; Wilson & Korn, 2007).

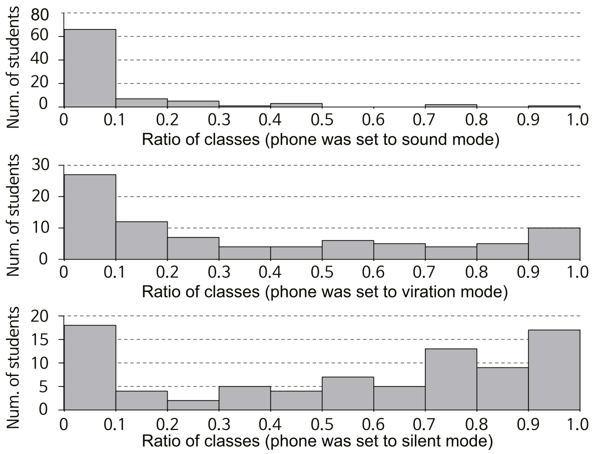

Fig. 7. Histograms of the fraction of the lectures in which a student's phone was set to a given mode.

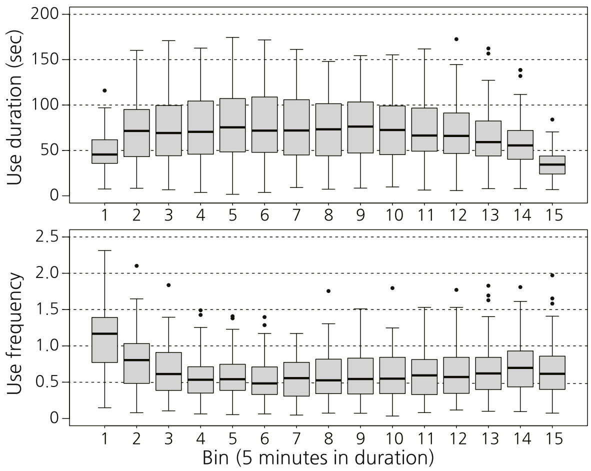

Fig. 8. Temporal rhythm of app use.

To quantify the temporal rhythm of app use we divided a lecture into 5-min bins. This bin size was typically used in prior research to measure students' attention level over time (Stuart & Rutherford, 1978; Wilson & Korn, 2007). A typical 90-min actual class length is approximately 75 min, because instructors end the class 10–15 min earlier to allow time for travel between classrooms (also confirmed by our participants). Thus, we have a total of 15 bins. In each bin, we calculated the mean use duration and frequency. For example, at the start of a given bin, a student used the phone for 1 min, and then used it for another minute after 1 min had passed. In this case, the use duration and frequency are given as 2 min and 2 times, respectively. It is possible that a session may span two consecutive bins. In the case of duration, we divide the session at the boundary time (multiple of 5 min) and added each segment into a separate bin. In the case of frequency, this session is only counted in the bin to which the start time belongs. For each student, we calculate two 15 dimensional vectors (i.e., duration and frequency) for each class and find the average values of those metrics.

Fig. 8 shows the use duration over time. Overall, the use duration per bin increases over time and plateaus around the second bin (10-min mark). It then starts decreasing around the twelfth bin (60-min mark). As shown in Fig. 8, the pattern of use frequency over time is very different from the pattern of use duration. Use frequency significantly drops in the first few bins and plateaus around the third bin (15-min mark), but it slightly increases towards the end. It is possible that at the beginning of a class, students may want to shortly use their phones before their classes start (e.g., a short checking of KakaoTalk or Facebook) or their usage may be interrupted by the instructor, which may result in more frequent, but shorter usage at the beginning of class. Likewise, frequent short checking happens towards the end of a class as well. Despite such transition behaviors, the plots indicate that students become distracted as early as when the class starts, and their distraction tendencies last until the class nears the end (around the 60-min mark) during which students start to regain their attention.

4.4. RQ4: predictors of in-class smartphone use

We identify what the predictors of in-class smartphone use are. Prior studies showed that the potential sources of distraction include the perceived difficulty of lecture materials and poor organization of lectures (Tesch et al., 2011), and students use mobile phones in particular when the class size is large (Tindell & Bohlander, 2012). For each class, we collected self-reports about the perceived difficulty levels every week and organizational efficiency of lectures in each month. We calculate the average values of the perceived difficulty and organizational efficiency. From the school registrars' database, we extracted the number of students enrolled. Additionally, we calculated each student's smartphone usage information: average daily use duration, use frequency, and average session duration to check how a student's *overall usage habit* is related to in-class smartphone usage. In addition, we used each student's SAS (Smartphone Addiction Scale) score (Kwon et al., 2013). Because each student typically takes 3–5 90-min classes, and in-class usage is largely influenced by each student's baseline usage behaviors, we perform multilevel regression analyses with an unconditional intercept model. We consider the following dependent variables: smartphone usage duration and frequency per class.

As shown in Table 3, the model for usage duration per class resulted in the adjusted R^2 value of 0.525 ($p < 0.001$), which is quite large. This model shows that class size and daily use/session duration are the key predictors of the in-class use duration. The perceived difficulty and lecture organization were only marginally significant. It appears that class size is considered as an important factor when it comes to smartphone use as shown in the prior survey study (Tindell & Bohlander, 2012). We also ran a similar regression analysis on the use frequency per class. The model's adjusted R^2 value was 0.525 ($p < 0.001$). Unlike the model for usage duration per class, the significant predictors were the average daily use frequency and the lecture organization (see the right column in Table 3).

Table 3
In-class smartphone usage models.

Independent variables	Use duration per class	Use frequency per class
	β	β
Avg. daily use duration	0.415***	−0.047
Avg. daily use frequency	0.092	0.631***
Avg. daily session duration	−0.320*	−0.019
Smartphone addiction score	0.036	−0.001
Perceived difficulty	0.083.	−0.025
Lecture organization	−0.078.	0.080*
Class size	0.106*	−0.016
Adjusted R^2	0.525***	0.525***

*$p < 0.05$, **$p < 0.01$, ***$p < 0.001$.

4.5. RQ5: phone use and semester grades

How are a student's smartphone usage behaviors related to their grades? For each student, we extracted two levels of phone usage metrics, namely day-level usage patterns and class-level usage patterns. At each level, we calculated the use duration and frequency. Furthermore, within each level, we considered the use patterns of major apps, namely Facebook, KakaoTalk, web browsers (Chrome, Samsung Browser, Naver), and the LMS app. We then performed hierarchical multiple regression on the student's mean semester grades. The first block considers a student's gender and smartphone addiction score, the second block additionally includes the student's day-level usage metrics, and the third block also includes the class-level usage metrics. As shown in Table 4, our initial model of gender and smartphone addiction scores was not significant (p = 0.100, adjusted R^2 = 0.032). However, our second model with day-level usage metrics was significant with the adjusted R^2 value of 0.149 (p = 0.021). Here, the daily use duration was the only significant predictor, with a negative correlation. It appears that how much a student uses a phone is more important than how frequently the student uses the phone. The third model with class-level usage metrics was significant with the adjusted R^2 value of 0.355 (p < 0.001), and the change of R^2 value was 0.206. As with the day-level model, the daily use duration was significant. We found that the daily web use duration was positively associated with grades. Interestingly, the use duration of web browsers during the class showed a very strong negative correlation with grades (p < 0.001). KakaoTalk use duration in the class was also negatively correlated. Interestingly, frequent usage of KakaoTalk and web browsers during the class was positively related to students' grades. It

Table 4
The models for students' grades.

Independent variables	Block 2 Day-level usage	Block 3 Class-level usage
	beta	beta
Gender	0.150	0.156
Smartphone addiction score	−0.078	0.040
Avg. daily use duration	−0.415***	−0.401***
Avg. daily use frequency	0.074	0.366
Avg. daily Facebook duration	−0.011	0.293
Avg. daily KakaoTalk duration	0.128	0.230
Avg. daily Web duration	0.014	0.559**
Avg. daily LMS duration	0.134	0.088
Avg. daily Facebook frequency	0.131	−0.025
Avg. daily KakaoTalk frequency	−0.099	−0.493
Avg. daily Web frequency	0.236	−0.306
Avg. daily LMS frequency	−0.203	−0.171
Avg. class use duration		−0.032
Avg. class use frequency		−0.364
Avg. class Facebook duration		−0.278
Avg. class KakaoTalk duration		−0.549*
Avg. class Web duration		−0.814***
Avg. class LMS duration		−0.052
Avg. class Facebook frequency		0.211
Avg. class KakaoTalk frequency		0.784*
Avg. class Web frequency		0.811**
Avg. class LMS frequency		0.108
Adjusted R^2	0.149**	0.355**
R^2 Change	0.146	0.206

*$p < 0.05$, **$p < 0.01$, ***$p < 0.001$.

seems like infrequent long usage appears to be more negative than frequent short usage in classrooms.

5. Discussion

We summarize the key findings and discuss the following: (1) prevalence of smartphone usage in classrooms, (2) reflection on attention duration in conventional learning contexts, (3) in-class smartphone use and its implications on academic performance, (4) methodological contributions to the fields, (5) practical implications on intervention design related to regulation policies and technical approaches, and (6) limitations and future work.

5.1. Prevalence of smartphone usage in classrooms

We collected smartphone usage and sensor data for 14 weeks from 84 college students and analyzed 196,704 h of smartphone usage data as well as 13,608 h of in-class smartphone usage data from 353 courses in classes of duration 90 min. Our results clearly showed that smartphone use in classrooms is prevalent as prior survey studies reported (McCoy, 2016). Increasingly more over time, younger generations have become more deeply wired into their mobile devices. We found that smartphone usage among college students has significantly increased over time: 5.78 h in 2017 vs. 3.75 h in 2012 (U. Lee et al., 2014).

Our in-class usage analyses showed that the students used their smartphones for 21.55 min per class on average. This means that they used their phones for more than a quarter of the class duration (28.1%), assuming effective class duration of 75 min for 90-min classes. Session analyses revealed a pattern that the students were distracted every 3–4 min for over a minute with their phones. Regression results showed that in-class usage was largely dependent on overall usage habits. It is likely that students developed a checking habit that involves brief and frequent content consumption and social interactions (e.g., Facebook and instant messaging) (Oulasvirta, Rattenbury, Ma, & Raita, 2012), which may have affected in-class usage as well. Their on- and off-task usage frequencies are on average 3.08 and 5.16, respectively. When the sum of on- and off-task usage frequencies is compared with the measured usage frequency, there was a significant *underestimation* (self-report: 8.25 vs. measurement: 12.28). In contrast, session duration per in-class use was *overestimated* (e.g., for off-tasks, self-report: 395.48 s vs. measurement: 117.52 s). Our measurement work provided empirical evidence of prevalent usage in classrooms as reported in a recent survey (McCoy, 2016): US students spent an average of 20.9% of class time for smartphone use, and the average use frequency was 11.43 per class session.

When considering the use duration/frequency of apps, our results showed that instant messaging, social media, and web browsers were the major apps that students used in classrooms. Owing to the quantitative nature of our work, it is difficult to judge what fraction of usage is related to off-tasks, because these apps could be also used for class purposes (e.g., sharing content and facilitating engagement) (Katz, 2005, pp. 305–317; Manca & Ranieri, 2013). However, we posit that class purpose usage would be relatively low because the prior studies showed the distractive nature of smartphones (I. Kim, Jung, Jung, Ko, & Lee, 2017; Tossell et al., 2014). Furthermore, it is less likely that smartphones have been used in classes for such prolonged duration over the entire semester, and personal instant messaging is rarely used for in-class activities. Our hierarchical regression analysis showed a negative relationship between excessive in-class phone use and grades. More in-depth log data analyses with experience sampling are required to justify this observation.

5.2. Reconsidering attention duration in classrooms

Contrary to our conventional beliefs that students' attention duration is about 10 min (despite recent debates about its validity (Bradbury, 2016)), we found empirical evidence that digital natives tended to easily redirect their attention to their phones. The students were distracted every 3–4 min for over a minute with their phones. Furthermore, the students' phone use started increasing as soon as the class started until it stabilized around the 15-min mark. Our findings suggest that we need to carefully consider digital technology use when estimating students' attention duration in classroom learning environments. In the past, students' attention was mostly estimated via proxy measures (e.g., degree of note-taking and recalling), external observations, and self-reports (Wilson & Korn, 2007). In addition, prior studies were mostly conducted in the 70s–90s (Stuart & Rutherford, 1978) where digital distractions would be far less problematic than in recent years. Currently, smartphone use has become deeply integrated into college students' everyday lives, and thus, we call for further studies on fine-grained attention modeling and measurement by systematically considering digital distractions.

5.3. In-class phone use and academic performance

Guided by prior studies in the field of education (Tesch et al., 2011; Tindell & Bohlander, 2012), we considered class parameters (i.e., perceived difficulty, class organization, and class size) and personal variables (i.e., daily usage time/frequency, session duration, and smartphone addiction score (Kwon et al., 2013)). Our multilevel regression analyses confirmed the importance of these parameters: the predictors of usage duration per class are daily usage duration and class size (with the marginal significance of perceived class difficulty and organization), and the predictors of usage frequency per class are daily usage frequency and perceived class organization. The *usage habit* strength (i.e., daily duration and frequency) is far stronger predictors than other variables. Unlike the prior belief (Samaha & Hawi, 2016), we did not find evidence suggesting that the level of self-reported smartphone addiction has a significant impact on in-class phone usage, which requires further investigation.

We performed a hierarchical multiple regression to study the predictive power of smartphone usage-related metrics for semester grades. None of the prior studies considered real in-class smartphone usage metrics. Our model confirmed that smartphone use is indeed closely related to grades ($R^2 = 0.343$). The major predictors include daily usage duration and various in-class usage of apps (i.e., instant messaging and web browser). Although the amount of Facebook usage is comparable to that of instant messaging and web browsing, it was not a statistically significant predictor. In fact, the impacts of Facebook usage on academic performance are also equivocal according to recent studies (Y. Wang & Mark, 2018; Michikyan et al., 2015). Interestingly, frequent usage of KakaoTalk and web browsers was *positively* associated with grades, although usage duration of those apps had a *negative* association. Despite overall negative concerns, these results hint that when students can self-regulate smartphone use well in classrooms (although it is challenging due to difficulties in boundary management (Lim et al., 2017)), brief and frequent smartphone use may have positive effects (e.g., supporting learning (Katz, 2005, pp. 305–317) or refreshing attention (Medina, 2014)).

5.4. Measurement methodological considerations

There has been a significant interest among researchers and practitioners in understanding personal technology use in classrooms. However, there has been a dearth of measurement research so far. Most prior studies rely on self-reports, which are subject to errors, and controlled experiments, with which it may be difficult to observe naturalistic usage behaviors. To our knowledge, the existing measurement studies did not specifically focus on in-class contexts although various sensing methods such as activity sensing and network traffic monitoring were used to understand student behaviors (R. Wang et al., 2014; Zhou et al., 2016). For example, the StudentLife project (R. Wang et al., 2014, 2015) mainly focused on passive sensing (i.e., activity, location, conversation, and sleep) without considering smartphone usage tracking due to methodological restrictions that most participants carried a secondary phone for sensing purposes. Passive Wi-Fi traffic monitoring helps in understanding overall student activities (Zhou et al., 2016), but detailed usage analysis is not feasible without deep packet inspection. Our work leveraged existing indoor localization and activity sensing techniques (Ashbrook & Starner, 2003; D. H.; Kim et al., 2010; Cho et al., 2015) for checking class attendance as well as detecting arrival/departure detection. As in the StudentLife project, we can incorporate passive sensing to enrich the analysis on in-class and out-of-class usage and activities. Our work provided a foundation for analyzing digital technology use behaviors in educational contexts by introducing novel data analysis opportunities to the fields of the educational data mining and learning analytics where smartphone usage and interaction data were not heavily considered. Our fine-grained analytics of personal technology use can be extended to intelligent tutoring or computer-supported learning environments (Romero & Ventura, 2010), in which we can provide feedback for supporting instructors and detecting undesirable student behaviors (e.g., low motivation, dropouts, cheating, failures, off-tasks).

5.5. Implications on intervention design

It is very important to properly deal with personal technologies in classrooms (Bayless, Clipson, & Wilson, 2013) by leveraging their potential as useful learning tools (Katz, 2005, pp. 305–317) but minimizing disruptive off-task usage (Chen & Yan, 2016; Levine, Waite, & Bowman, 2012). Consistent with prior results with in-class laptop use (Fried, 2008; Sana et al., 2013), our results indicate that in-class off-task phone usage such as instant messaging and web browsing should be properly regulated in order to minimize negative impacts on learning. We recommend researchers and practitioners to consider our results when designing *in-class technology use policies*, developing *digital literacy programs*, and adopting *technical solutions*.

Prior studies emphasized the importance of properly setting in-class technology use policies in the syllabus and their consistent enforcement (Hopke & Marsh, 2011; Tindell & Bohlander, 2012). Our results suggest that the instructors should proactively engage in regulating students' off-task app use (e.g., frequent instant messaging and excessive web browsing) as the prior studies recommended (Bayless et al., 2013; Tindell & Bohlander, 2012). In addition, a prior study documented that college students as digital natives can easily adopt unfamiliar technologies through digital literacy education (Ng, 2012). Thus, digital literacy programs of teaching students how to regulate their use of personal technology can also be beneficial in controlling in-class smartphone use.

Our results help in designing various technical approaches for mitigating distractive smartphone use in classrooms. Visualizing fine-grained interactions, as well as attention fragmentation, will help students to better reflect upon their in-class multitasking behaviors. Furthermore, we can provide proactive support features optimized for in-class intervention such as context-aware reminders (I. Kim et al., 2017), automatic ringer-mode re-configuration (Siewiorek et al., 2003), and automatic rescheduling of notification delivery by identifying opportune moments (e.g., in-class breakpoints) (Horvitz, Koch, & Apacible, 2004; Okoshi, Tsubouchi, Taji, Ichikawa, & Tokuda, 2017). Universities' learning management systems can include usage tracking and context-aware features (Lochtefeld, Bohmer, & Ganev, 2013); for example, a student's location and class schedule can be used to provide a proactive warning or nudging on smartphone overuse during classes (Okeke, Sobolev, Dell, & Estrin, 2018). In this case, a design requirement is to ensure autonomy and agency (e.g., voluntary participation, power distance mitigation (Rotman, 2010)) as well as privacy preservation (e.g., data privacy and sensitivity).

5.6. Limitations and future work

As the study was performed at a single site, further work is needed to improve the generalizability of our findings. Our work only considered freshmen in order to control the effects of class rank, owing to recent findings that students with higher rank tend to show different usage behaviors (Junco, 2015; Pempek, Yermolayeva, & Calvert, 2009). There should be further measurement studies with diverse class ranks possibly across different sites and cultures (because there could be some cultural influences on phone use such as smartphone dependence and normative behaviors in classrooms) (S.-H. Lee, 2015). Our work mainly focused on the quantitative understanding of usage patterns in classrooms, unfortunately missing detailed accounts of usage behaviors that qualitative methods could have picked up. Our analytic approach is limited in that we cannot differentiate on- and off-task usage, despite ample evidence of off-task usage (as predictors of grade estimation). Further work is needed to clarify on- and off-task app use in classrooms. This will help us to train a machine learning model for automatic identification of off-task app use as in prior studies on intelligent tutoring systems (Baker, 2007). Our regression model mainly considered smartphone usage metrics without considering other factors that may influence usage behaviors. As prior studies suggested (Mark et al., 2014; R.; Wang et al., 2014; Baker, D'Mello, Rodrigo, & Graesser, 2010), for future work, we could consider personal traits, health conditions, affective states, and activity metrics to build more sophisticated multitasking models. Additionally, individuals' behavioral trends (R. Wang et al., 2015) and circadian rhythms (or chronotypes) (Murnane et al., 2016) could be incorporated.

6. Conclusions

We conducted a long-term measurement study by analyzing 196,704 h of smartphone usage data for 14 weeks as well as class-specific self-reported data of 445 classes over the semester collected from 84 college students in Korea. We found that students used their phones for more than 20 min per class, which is over 25% of the effective class duration. Phone usage was prevalent throughout the class, and phone distractions occurred every 3–4 min for over a minute in duration. Students significantly underestimated in-class phone usage frequency (self-report: 8.25 vs. logged measurement: 12.28). Predictors of in-class usage included daily usage behaviors, class size, and lecture organization. Phone usage patterns such as daily and in-class use habits had a negative relationship with student grades.

Personal technology use among college students is prevalent. Our students used to be mobile pioneers but are now digital natives (known as the mobile-first generation or Generation Z). Smartphone usage is deeply integrated into students' everyday lives. Our results clearly demonstrated that smartphone usage is indeed widespread in college classrooms, and despite potential benefits of smartphones as learning and social support tools, excessive phone usage can be considered harmful to academic performance. We call for further studies on student modeling, educational decision making, and pedagogical support related to in-class smartphone usage.

Acknowledgements

This work was supported by LG Yonam Foundation and Next-Generation Information Computing Development Program through the National Research Foundation of Korea (NRF) funded by the Ministry of Science and ICT (NRF-2017M3C4A7083534).

References

Adler, R. F., & Benbunan-Fich, R. (2013). Self-interruptions in discretionary multitasking. *Computers in Human Behavior, 29*(4), 1441–1449.
Ames, M. G. (2013). Managing mobile multitasking: The culture of iphones on Stanford campus. *Proceedings of the 2013 conference on computer supported cooperative work* (pp. 1487–1498). .
Ashbrook, D., & Starner, T. (2003). Using gps to learn significant locations and predict movement across multiple users. *Personal and Ubiquitous Computing, 7*(5), 275–286.
Bahl, P., & Padmanabhan, V. N. (2000). Radar: An in-building rf-based user location and tracking system. *Infocom 2000. nineteenth annual joint conference of the ieee computer and communications societies. proceedings. ieee: Vol. 2*, (pp. 775–784).
Baker, R. S. (2007). Modeling and understanding students' off-task behavior in intelligent tutoring systems. *Proceedings of the sigchi conference on human factors in computing systems* (pp. 10591068). .
Baker, R. S., D'Mello, S. K., Rodrigo, M. M. T., & Graesser, A. C. (2010). Better to be frustrated than bored: The incidence, persistence, and impact of learners' cognitive-affective states during interactions with three different computer-based learning environments. *International Journal of Human-Computer Studies, 68*(4), 223–241.
Bayless, M. L., Clipson, T. W., & Wilson, S. (2013). *Faculty perceptions and policies of students' use of personal technology in the classroom*.
Ben Abdesslem, F., Phillips, A., & Henderson, T. (2009). Less is more: Energy-efficient mobile sensing with senseless. *Proceedings of the 1st acm workshop on networking, systems, and applications for mobile handhelds* (pp. 61–62). New York, NY, USA: ACM. Retrieved from http://doi.acm.org/10.1145/1592606.1592621 doi: 10.1145/1592606.1592621.
Böhmer, M., Hecht, B., Schöning, J., Krüger, A., & Bauer, G. (2011). Falling asleep with angry birds, facebook and kindle: A large scale study on mobile application usage. *Proceedings of the 13th international conference on human computer interaction with mobile devices and services* (pp. 4756). .
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.
Bradbury, N. A. (2016). Attention span during lectures: 8 seconds, 10 minutes, or more? *Advances in Physiology Education, 40*(4), 509–513.
Chen, Q., & Yan, Z. (2016). Does multitasking with mobile phones affect learning? A review. *Computers in Human Behavior, 54*, 34–42.
Cho, D.-K., Lee, U., Noh, Y., Park, T., & Song, J. (2015). Placewalker: An energy-efficient place logging method that considers kinematics of normal human walking. *Pervasive and Mobile Computing, 19*, 24–36.
Czerwinski, M., Horvitz, E., & Wilhite, S. (2004). A diary study of task switching and interruptions. *Proceedings of the sigchi conference on human factors in computing systems* (pp. 175–182). .
David, P., Kim, J.-H., Brickman, J. S., Ran, W., & Curtis, C. M. (2015). Mobile phone distraction while studying. *New Media & Society, 17*(10), 1661–1679.
Falaki, H., Mahajan, R., Kandula, S., Lymberopoulos, D., Govindan, R., & Estrin, D. (2010). Diversity in smartphone usage. *Proceedings of the 8th international conference on mobile systems, applications, and services* (pp. 179–194). New York, NY, USA: ACM.
Ferreira, D., Goncalves, J., Kostakos, V., Barkhuus, L., & Dey, A. K. (2014). Contextual experience sampling of mobile application micro-usage. *Proceedings of the 16th international conference on human-computer interaction with mobile devices & services* (pp. 91–100). New York, NY, USA: ACM.
Fried, C. B. (2008). In-class laptop use and its effects on student learning. *Computers & Education, 50*(3), 906–914.
Gaudreau, P., Miranda, D., & Gareau, A. (2014). Canadian university students in wireless classrooms: What do they do on their laptops and does it really matter? *Computers & Education, 70*, 245–255.
Gingerich, A. C., & Lineweaver, T. T. (2014). Omg! texting in class = u fail:(empirical evidence that text messaging during class disrupts comprehension. *Teaching of Psychology, 41*(1), 44–51.
González, V. M., & Mark, G. (2004). Constant, constant, multi-tasking craziness: Managing multiple working spheres. *Proceedings of the sigchi conference on human*

factors in computing systems (pp. 113–120). .
Hopke, K. D., & Marsh, P. A. (2011). Student cell phone use in college classrooms. *Psychology and Education, 48*(1), 47–58.
Horvitz, E., Koch, P., & Apacible, J. (2004). Busybody: Creating and fielding personalized models of the cost of interruption. *Proceedings of the 2004 acm conference on computer supported cooperative work* (pp. 507–510). .
Jin, J., & Dabbish, L. A. (2009). Self-interruption on the computer: A typology of discretionary task interleaving. *Proceedings of the sigchi conference on human factors in computing systems* (pp. 1799–1808). .
Junco, R. (2015). Student class standing, facebook use, and academic performance. *Journal of Applied Developmental Psychology, 36*, 18–29.
Junco, R., & Cotten, S. R. (2012). No a 4 u: The relationship between multitasking and academic performance. *Computers & Education, 59*(2), 505–514.
Katz, J. E. (2005). *Mobile phones in educational settings. A sense of place: The global and the local in mobile communication*.
Kim, I., Jung, G., Jung, H., Ko, M., & Lee, U. (2017). Let's focus: Mitigating mobile phone use in college classrooms. *Proceedings of the ACM on Interactive, Mobile, Wearable and Ubiquitous Technologies, 1*(3), 63.
Kim, D. H., Kim, Y., Estrin, D., & Srivastava, M. B. (2010). Sensloc: Sensing everyday places and paths using less energy. *Proceedings of the 8th acm conference on embedded networked sensor systems* (pp. 43–56). .
Kolek, E. A., & Saunders, D. (2008). Online disclosure: An empirical examination of undergraduate facebook profiles. *NASPA Journal, 45*(1), 1–25.
Kuznekoff, J. H., & Titsworth, S. (2013). The impact of mobile phone usage on student learning. *Communication Education, 62*(3), 233–252.
Kwon, M., Lee, J.-Y., Won, W.-Y., Park, J.-W., Min, J.-A., Hahn, C., ... Kim, D.-J. (2013). Development and validation of a smartphone addiction scale (sas). *PLoS One, 8*(2), e56936.
Lee, S.-H. (2015). Mobile phone culture: The impacts of mobile phone use. *Encyclopedia of mobile phone behavior* (pp. 658–672). IGI Global.
Lee, U., Lee, J., Ko, M., Lee, C., Kim, Y., Yang, S., ... Song, J. (2014). Hooked on smartphones: An exploratory study on smartphone overuse among college students. *Proceedings of the 32nd annual acm conference on human factors in computing systems* (pp. 2327–2336). .
Lepp, A., Barkley, J. E., & Karpinski, A. C. (2015). The relationship between cell phone use and academic performance in a sample of us college students. *Sage Open, 5*(1) 2158244015573169.
Levine, L. E., Waite, B. M., & Bowman, L. L. (2012). Mobile media use, multitasking and distractibility. *International Journal of Cyber Behavior, Psychology and Learning, 2*(3), 15–29.
Lim, H., Arawjo, I., Xie, Y., Khojasteh, N., & Fussell, S. R. (2017). Distraction or life saver?: The role of technology in undergraduate students' boundary management strategies. *Proceedings of the ACM on Human-Computer Interaction, 1*, 68 CSCW.
Lochtefeld, M., Bohmer, M., & Ganev, L. (2013). Appdetox: Helping users with mobile app addiction. *Proceedings of the international conference on mobile and ubiquitous multimedia* (pp. 43). .
Manca, S., & Ranieri, M. (2013). Is it a tool suitable for learning? A critical review of the literature on facebook as a technology-enhanced learning environment. *Journal of Computer Assisted Learning, 29*(6), 487–504.
Mark, G., González, V. M., & Harris, J. (2005). No task left behind?: Examining the nature of fragmented work. *Proceedings of the sigchi conference on human factors in computing systems* (pp. 321–330). .
Mark, G., Wang, Y., & Niiya, M. (2014). Stress and multitasking in everyday college life: An empirical study of online activity. *Proceedings of the sigchi conference on human factors in computing systems* (pp. 41–50). .
Mayer, R. E., & Moreno, R. (2003). Nine ways to reduce cognitive load in multimedia learning. *Educational Psychologist, 38*(1), 43–52.
McCoy, B. R. (2016). *Digital distractions in the classroom phase ii: Student classroom use of digital devices for non-class related purposes*.
Medina, J. (2014). *Brain rules: 12 principles for surviving and thriving at work, home, and school.* Pear Press.
Michikyan, M., Subrahmanyam, K., & Dennis, J. (2015). Facebook use and academic performance among college students: A mixed-methods study with a multi-ethnic sample. *Computers in Human Behavior, 45*, 265–272.
Murnane, E. L., Abdullah, S., Matthews, M., Kay, M., Kientz, J. A., Choudhury, T., ... Cosley, D. (2016). Mobile manifestations of alertness: Connecting biological rhythms with patterns of smartphone app use. *Proceedings of the 18th international conference on human-computer interaction with mobile devices and services* (pp. 465–477). New York, NY, USA: ACM.
Ng, W. (2012). Can we teach digital natives digital literacy? *Computers & Education, 59*(3), 1065–1078.
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* (pp. 4:1–4:12). New York, NY, USA: ACM.
Okoshi, T., Tsubouchi, K., Taji, M., Ichikawa, T., & Tokuda, H. (2017). Attention and engagement- awareness in the wild: A large-scale study with adaptive notifications. *2017 ieee international conference on pervasive computing and communications (percom)* (pp. 100–110). .
Ophir, E., Nass, C., & Wagner, A. D. (2009). Cognitive control in media multitaskers. *Proceedings of the National Academy of Sciences, 106*(37), 15583–15587.
Oulasvirta, A., Rattenbury, T., Ma, L., & Raita, E. (2012). January). Habits make smartphone use more pervasive. *Personal and Ubiquitous Computing, 16*(1), 105–114. Retrieved from https://doi.org/10.1007/s00779-011-0412-2 doi: 10.1007/s00779-011-0412-2.
Paek, J., Kim, J., & Govindan, R. (2010). Energy-efficient rate-adaptive gps-based positioning for smartphones. *Proceedings of the 8th international conference on mobile systems, applications, and services* (pp. 299–314). New York, NY, USA: ACM. https://doi.org/10.1145/1814433.1814463. Retrieved from http://doi.acm.org/10.1145/1814433.1814463.
Pasek, J., More, E., & Hargittai, E. (2009). Facebook and academic performance: Reconciling a media sensation with data. *First Monday, 14*(5)https://firstmonday.org/article/view/2498/2181.
Pempek, T. A., Yermolayeva, Y. A., & Calvert, S. L. (2009). College students' social networking experiences on facebook. *Journal of Applied Developmental Psychology, 30*(3), 227–238.
Ravizza, S., Uitvlugt, M., & Fenn, K. (2017). The negative effects of laptop internet use during class. *Neuroscience Letters, 637*, 44–49.
Romero, C., & Ventura, S. (2010). Educational data mining: A review of the state of the art. *IEEE Transactions on Systems, Man, and Cybernetics, Part C (Applications and Reviews), 40*(6), 601618.
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.
Rotman, D. (2010). Constant connectivity, selective participation: Mobile-social interaction of students and faculty. *Chi'10 extended abstracts on human factors in computing systems* (pp. 4333–4338). .
Samaha, M., & Hawi, N. S. (2016). Relationships among smartphone addiction, stress, academic performance, and satisfaction with life. *Computers in Human Behavior, 57*, 321–325.
Sana, F., Weston, T., & Cepeda, N. J. (2013). Laptop multitasking hinders classroom learning for both users and nearby peers. *Computers & Education, 62*, 24–31.
Siewiorek, D., Smailagic, A., Furukawa, J., Krause, A., Moraveji, N., Reiger, K., ... Wong, F. L. (2003). *Sensay: A context-aware mobile phone* (pp. 248). null.
Stuart, J., & Rutherford, R. (1978). Medical student concentration during lectures. *The Lancet, 312*(8088), 514–516.
Tesch, F., Coelho, D., & Drozdenko, R. (2011). We have met the enemy and he is us: Relative potencies of classroom distractions. *Business Education Innovation Journal, 3*(2).
Tindell, D. R., & Bohlander, R. W. (2012). The use and abuse of cell phones and text messaging in the classroom: A survey of college students. *College Teaching, 60*(1), 1–9.
Tossell, C. C., Kortum, P., Shepard, C., Rahmati, A., & Zhong, L. (2014). You can lead a horse to water but you cannot make him learn: Smartphone use in higher education. *British Journal of Educational Technology, 46*(4), 713–724.
Waite, B. M., Lindberg, R., Ernst, B., Bowman, L. L., & Levine, L. E. (2018). Off-task multitasking, note-taking and lower-and higher-order classroom learning. *Computers & Education, 120*, 98–111.
Wang, R., Chen, F., Chen, Z., Li, T., Harari, G., Tignor, S., ... Campbell, A. T. (2014). Studentlife: Assessing mental health, academic performance and behavioral trends of college students using smartphones. *Proceedings of the 2014 acm international joint conference on pervasive and ubiquitous computing* (pp. 3–14). .

Wang, R., Harari, G., Hao, P., Zhou, X., & Campbell, A. T. (2015). Smartgpa: How smartphones can assess and predict academic performance of college students. *Proceedings of the 2015 acm international joint conference on pervasive and ubiquitous computing* (pp. 295–306). .

Wang, Y., & Mark, G. (2018). The context of college students' facebook use and academic performance: An empirical study. *Proceedings of the 2018 chi conference on human factors in computing systems* (pp. 418). .

Wang, Y., Niiya, M., Mark, G., Reich, S. M., & Warschauer, M. (2015). Coming of age (digitally): An ecological view of social media use among college students. *Proceedings of the 18th acm conference on computer supported cooperative work & social computing* (pp. 571–582). .

Wei, F.-Y. F., & Wang, Y. K. (2010). Students' silent messages: Can teacher verbal and nonverbal immediacy moderate student use of text messaging in class? *Communication Education, 59*(4), 475–496.

Wickens, C. D., Hollands, J. G., Banbury, S., & Parasuraman, R. (2015). *Engineering psychology & human performance.* Psychology Press.

Wilmer, H. H., Sherman, L. E., & Chein, J. M. (2017). Smartphones and cognition: A review of research exploring the links between mobile technology habits and cognitive functioning. *Frontiers in Psychology, 8*, 605.

Wilson, K., & Korn, J. H. (2007). Attention during lectures: Beyond ten minutes. *Teaching of Psychology, 34*(2), 85–89.

Wohn, D. Y., & Birnholtz, J. (2015). From ambient to adaptation: Interpersonal attention management among young adults. *Proceedings of the 17th international conference on human-computer interaction with mobile devices and services* (pp. 26–35). .

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 & Education, 58*(1), 365–374.

Zhou, M., Ma, M., Zhang, Y., SuiA, K., Pei, D., & Moscibroda, T. (2016). Edum: Classroom education measurements via large-scale wifi networks. *Proceedings of the 2016 acm international joint conference on pervasive and ubiquitous computing* (pp. 316–327). .