Hooked on Smartphones: An Exploratory Study on Smartphone Overuse among College Students

Uichin Lee, Joonwon Lee, Minsam Ko, Changhun Lee†, Yuhwan Kim§, Subin Yang
Koji Yatani‡, Gahgene Gweon, Kyong-Mee Chung♯, Junehwa Song§
Knowledge Service Engineering Department, KAIST, †Web Science and Technology Division, KAIST
‡Microsoft Research Asia, §Computer Science Department, KAIST
♯Department of Psychology, Yeonsei University
{uclee, joonwonlee, msko, changhun, yuhwan, subin.yang, ggweon, junesong}@kaist.ac.kr
‡koji@microsoft.com ♯kmchung@yonsei.ac.kr

ABSTRACT
The negative aspects of smartphone overuse on young adults, such as sleep deprivation and attention deficits, are being increasingly recognized recently. This emerging issue motivated us to analyze the usage patterns related to smartphone overuse. We investigate smartphone usage for 95 college students using surveys, logged data, and interviews. We first divide the participants into risk and non-risk groups based on self-reported rating scale for smartphone overuse. We then analyze the usage data to identify between-group usage differences, which ranged from the overall usage patterns to app-specific usage patterns. Compared with the non-risk group, our results show that the risk group has longer usage time per day and different diurnal usage patterns. Also, the risk group users are more susceptible to push notifications, and tend to consume more online content. We characterize the overall relationship between usage features and smartphone overuse using analytic modeling and provide detailed illustrations of problematic usage behaviors based on interview data.

Author Keywords
Smartphone Overuse; Measurement

ACM Classification Keywords
H.5.m. Information Interfaces and Presentation (e.g. HCI): Miscellaneous

INTRODUCTION
The popularity of smartphones has been increasing rapidly in recent years. In most of developed countries, the rate of smartphone adoption exceeded 50% in the first half of 2012 [2]. Thus, smartphones have now become an integral part of the daily lives of many individuals. However, negative aspects of their use have emerged, such as the disruption of social interactions. In addition, researchers have found close relationship between their overuse and poor mental health (e.g., sleep deprivation and attention deficits) [24]. Interactive characteristics of smartphones contain inducing and reinforcing features that promote excessive usage behaviors [17, 10]. For example, Oulasvirta et al. [25] have demonstrated that frequently checking dynamic content (e.g., updates from online social networks) on mobile devices weakens self-regulation, which may lead to smartphone overuse [17, 10].

Researchers have tried to identify the problematic usage patterns that are related with smartphone overuse [5, 16], mainly through self-reported scale which are subjective and prone to recall errors [27]. Thus far, relatively little information is known about how smartphone overuse is reflected in the actual use of smartphones. In this paper, we investigate the college students as a focusing lens on emerging technological issues. Smartphones are used widely by college students. In the first half of 2012, the smartphone adoption rates in the UK and South Korea were 72% and 86%, respectively [2]. Moreover, college students are considered to be vulnerable to technology overuse because of their developmental dynamics and relative independence from social roles and expectations [14].

We collected actual smartphone usage data from 95 college students (over 50,000 hours of usage data) and performed an exploratory data analysis along with the smartphone addiction psychometric scale [15]. Our smartphone usage logger allowed for unobtrusive monitoring that has minimal impacts on user behavior [27]. Participants were divided into two groups based on rating score: a risk group (whose scores indicated a potential for smartphone overuse) and a non-risk group. We examined the differences in the usage patterns of the two groups by comparing various usage features, which were extracted from the dataset (e.g., overall usage and content-specific usage). These usage differences were validated by analytic modeling, and also corroborated by the findings from our interviews.

We identified several usage patterns that were closely related to smartphone overuse. The risk group spent more time on smartphone use per day (risk group: 253.0 min *vs.* non-risk group: 207.0 min) compared with the non-risk group,

CHI 2014, April 26–May 1, 2014, Toronto, Ontario, Canada.

ACM 978-1-4503-2473-1/14/04 ...$15.00.
http://dx.doi.org/10.1145/2556288.2557366

and there were also diurnal usage differences (greater usage in the morning and evening). The usage sessions that were initiated by the push notifications were longer for the risk group, which demonstrated that notifications acted as external cues related to problematic usage patterns. The risk group consumed significantly more online content that can provide instant gratifications (e.g., pastimes, entertainment, and information seeking).

These findings were supported by the results of our analytic modeling and the analysis of our interview data. Regression analysis showed that the key components of smartphone overuse, such as interference, withdrawal, and tolerance, were closely related to usage features. In addition, we used machine learning techniques to test the predictive power of the usage features and found that smartphone overuse could be classified accurately based on these features with an F-score of 0.87. The interview results provided more detailed examples of problematic usage behavior, such as limited self-control when consuming online content (e.g., aimlessly following web/Facebook links while in bed). Our key findings support the existing theories related to technological addictions [11, 17, 10]. In particular, repeated consumption of online content may lead to addictive behaviors, and problematic behaviors depend mainly on specific functions rather than the volume of usage.

BACKGROUND AND RELATED WORK

Technological Addiction and Smartphone Overuse

Technological addictions are defined as behavioral (non-chemical) addictions, where interactive components of computer devices can have inducing and reinforcing features that may promote addictive tendencies (e.g., tolerance, withdrawal, interference, and relapse) [11]. Despite recent clinical and scientific evidence, however, there is a lack of agreement on the existence of technological addictions. According to Morahan-Martin's report [22], some researchers argue that the pathological use of technology services is driven by attempts to avoid the underlying mental or social problems. Others contend that technology is merely a medium for service delivery, and that obsessive use is related to specific services such as gambling and games. These arguments are debatable, but an awareness of technological addictions is growing in scientific communities. The Diagnostic and Statistical Manual of Mental Disorders, Fifth Edition (DSM-5), which was released in May 2013 by the American Psychiatric Association (APA), officially recognized behavioral addictions for the first time and recommended further research into existing technological addictions for later inclusion [1].

Prior studies have identified contributing factors that make digital media, such as the Internet and cell phones, attractive and often addictive [11, 17, 10]. Digital media provide easy and convenient ways of accessing a large amount of online content (e.g., music, news, and games) and maintaining social relationships. Such access gives instant gratifications to users (e.g., interpersonal utility, pastimes, information seeking, and entertainment), in turn reinforcing continuous usage of it [17, 10]. This concept is known as operant conditioning, which is one of the fundamental learning theories in behavioral psychology [21]. Excessive usage may follow subsequently, possibly due to deficient self-regulation or maladaptive thinking [11, 18, 17, 10].

Previous studies of Internet addiction showed that excessive use of online communication and games occurs often, which is related to various psychological factors, including social anxiety, depression, impulsivity, self-esteem/identity deficits, and situational stress during life changing events [23, 28, 19]. Similarly, excessive usage has also been studied in the context of mobile phone use [5, 13]. Carbonell et al. [5] reported that the excessive use of mobile phones by teenagers is attributable to text-messaging and mobile games. Hwang et al. [13] showed that it is also related with various psychological factors such as social anxiety, depression, and impulsivity. In contrast to previous studies [5, 13], we perform an exploratory data analysis of real usage datasets to uncover the usage features related to smartphone overuse, and validate the differences between usage patterns using analytic modeling and analysis of interview data.

Smartphone Usage Studies

Cui and Roto [6] found that the main use of smartphones was task-oriented with goals of information seeking, communications, online transactions, and managing personal information. In an observational study of smartphone usage on the Stanford campus, Ames [3] showed that the availability of always-on connectivity meant that the students had to exhibit the techno-social practices of balancing their extended networks with the immediate surroundings and to limit the negative impacts of smartphone usage (e.g., social pressure, and multi-tasking). Harmon and Mazmanian [12] identified two themes of smartphone use that are reflected in commercials, where one theme recommends the deep integration of smartphones in daily life, and the other urges people toward disintegration. Oulasvirta et al. [25] reported that the use of mobile devices may lead to the development of a checking habit that involves brief and frequent content consumption (e.g., checking emails and Facebook updates).

A few studies have characterized how users access their mobile devices during their everyday lives. For example, Falaki et al. [8] studied the usage patterns of Android and Windows Mobile phones by analyzing user interactions, app use, network traffic, and energy consumption. Despite the differences in the objectives and frequencies of smartphone use, they found that there were similarities in the fine-grained usage patterns of users (e.g., the session time distribution). Böhmer et al. [4] analyzed a large-scale, global dataset primarily from the US and Europe. They found that users typically spent almost one hour per day on smartphones, and that the average session duration was less than one minute. Their study also indicated the different time dependencies for app usage. For example, news apps were accessed most frequently in the morning, whereas communication apps (email and SMS) were used throughout the day. Although earlier studies [8, 4] provided general overviews of smartphone usage, they did not investigate the usage patterns related to smartphone overuse. By contrast, our study examines the similarities and differences between the smartphone usage among users with overuse risks and those without.

HCI Research into Addictive Behavior

Addiction-related studies have been increasing recently in the HCI community. A major goal of these studies is to explore the main factors to develop effective addiction intervention mechanisms. Seay and Kraut [26] showed that self-regulation is critical for controlling online gaming behaviors, and they considered how it can be incorporated into the game designs to prevent addictive behaviors. Another direction is to design new computing services or to simply use existing services to mitigate problematic use and assist traditional treatments. Wang et al. [29] built a smartphone-based system for tracking sobriety after patients left a rehabilitation center and for communicating relevant information to fellow patients and counselors. Yarosh [30] performed an observational study on the role of technology in alcoholism recovery. Our study attempts to identify the usage patterns related to smartphone overuse and to provide several guidelines to facilitate the design of intervention software.

METHODOLOGY

Participants

We investigated the smartphone usage behaviors of undergraduate students at a large university in Korea during the fall semester of 2012. We selected this homogeneous user population to help improve the internal validity of this study. This is a common practice in psychological studies [5, 28]. Selecting college students as the study groups is also beneficial because the smartphone adoption rate among young adults is very high [2]. In addition, they are vulnerable to technology overuse due to their developmental dynamics (e.g., identity formation) and relative independence from social roles and expectations [14]. We recruited students in 2012 during the time period between the mid-term and final exams in order to avoid the influence of exams affecting our study. Ninety-five students participated in the experiment for at least two weeks. The average age of the participants was 20.6 years (SD 1.7); and 67 (i.e., 70.5%) were male. Data were collected for an average of 26.8 days (SD 9.5). All of the participants were compensated for their participation with a gift that was equivalent to $20.

Smartphone Usage Logging

We developed the *SmartLogger* software to log a variety of application events (active/inactive apps, touch and text input events, web browsing URLs, and notification events), system events (power on/off and screen on/off/unlock), and phone events (calls and SMS). SmartLogger operates as an Android accessibility service. After an accessibility service has been enabled in the system settings, it runs automatically in the background.

Our approach can be considered as an unobtrusive method because the participants were generally unaware that they were being observed, which meant that their usage behavior was not affected by the data collection process [27]. We examined whether there were any significant differences in the daily usage levels during the experimental period (excluding the data obtained on the day of installation). We found that there were statistically significant differences only in the first two days compared with the remainder of the first week, and there were

F1	*"My school grades (or work productivity) dropped due to excessive smartphone use."* *"People often complained about excessive smartphone use."*
F2	*"Using a smartphone is more enjoyable than spending time with my family or friends."* *"When I cannot use my smartphone, I feel like I have lost the entire world."*
F3	*"It would be distressing if I am not allowed to use my smartphone."* *"I become restless and nervous when smartphone use is impeded."*
F4	*"Even when I think I should stop, I continue to use my smartphone."* *"Spending a lot of time on my smartphone has become a habit."*

Table 1. Illustration of Smartphone Addiction Proneness Scale (its subfactors include F1: Interference, F2: Virtual World, F3: Withdrawal, and F4: Tolerance)

no significant changes subsequently. Therefore, we excluded the data recorded during the first two days of participation from the subsequent analysis.

User Surveys and Interviews

Before initiating the data collection process, we administered surveys to acquire demographic data and measure the level of smartphone overuse. We used the *Smartphone Addiction Proneness Scale for Adults* [15], which is an established scale for smartphone overuse. The scale comprises 15 four-point Likert-scale questions (see the examples in Table 1), which measure four factors associated with addictive behaviors (interference, virtual world, withdrawal, and tolerance). The scale was designed to classify users into three groups (i.e., high-risk, at-risk, and normal user groups) [15]. In our study, we divided the participants into two groups: i.e., high/at-risk (total score $\geq$ 40 or F1: interference score $\geq$ 14) and non-risk (the remaining participants). This decision was based on the fact that the sample size would have been too small for subsequent analyses otherwise ($<$ 3% as the threshold used to separate the high-risk group was set at 2 SD above the mean in the scale). Figure 1 shows the distribution of the participants based on the F1 (interference) score and the total score. The risk and non-risk groups comprised 36 and 59 participants, respectively.

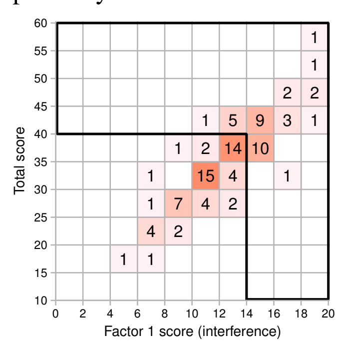

Figure 1. Score distribution of the experimental participants (high/at-risk groups: factor 1 score $\geq$ 14 or total score $\geq$ 40, non-risk group: remaining participants)

After the data collection process, we performed additional surveys and interviews to supplement the results from our quantitative data analysis and to characterize problematic usage behaviors. In the exit-surveys, we explored the usage of mobile instant messaging and its negative impacts. In addition, we selected seven participants (four risk and three non-risk) to participate in semi-structured interviews, where our selection criterion required that their average smartphone usage per day exceed four hours. We aimed to enhance our understanding of the usage characteristics determined in the quantitative analysis and to identify more detailed themes related to problematic usage behaviors. All of the interviews were transcribed and open coding was performed to facilitate content analysis.

Data Analysis Model

Active smartphone usage can be represented based on a series of sessions, as shown in Figure 2. A session was defined as the time between turning on and off events of the screen.

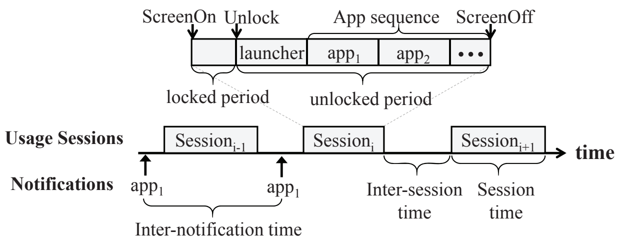

Figure 2. Usage data analysis model

During a typical session, a participant first unlocked the phone and then operated a series of apps; we excluded launcher usage in between apps. Several types of events, such as notifications and battery charging, could switch on the screen automatically. In our analysis, we filtered out the events that did not elicit further user interactions (i.e., screen unlocking). Users may have apps that provided notification messages (e.g., for new chats and emails). In Android, an icon appears in the notification area, and users can open the notification drawer to check the details and interact with their apps. For app-specific usage analysis, we designed simplified categories based on existing app categories from app stores and a previous study [4].

OVERALL DIFFERENCES IN USAGE PATTERNS

We analyzed the aggregated usage, session-level usage, and temporal usage patterns to identify differences in the specific usage levels of the risk and non-risk groups. The overall usage results are presented in Figure 3 and Figure 4.

Aggregated Usage: For a given day, the daily usage time was defined as the sum of all the session times. We calculated the mean daily usage time for each participant. There was a significant difference in the daily usage time, where the risk group (253.0 min, SD: 90.9, p = .011, Cohen's d = 0.54) had a longer average smartphone usage period than the non-risk group (207.4 min, SD: 77.2). We also measured how often the participants interacted with their smartphones by calculating the mean session frequency per day and the mean inter-session duration (i.e., the interval between two consecutive sessions), but there were no significant differences. The risk group had a slightly higher mean session frequency (risk: 111.5 *vs.* non-risk: 100.1, p = .146, Cohen's d = 0.31) and a shorter mean inter-session time (risk: 729.1 s *vs.* 816.6 s, p = .216, Cohen's d = 0.26), but the differences were not significant.

Session-level Usage: We also investigated whether the usage patterns within a session differed between the groups. We

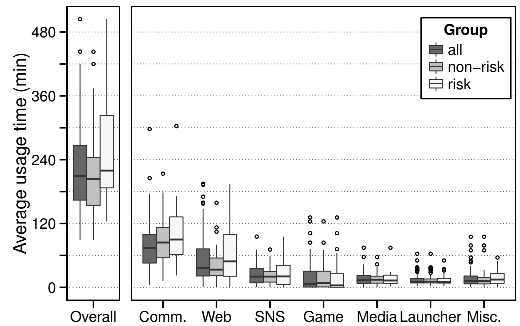

Figure 3. Usage amount: overall and app-specific results

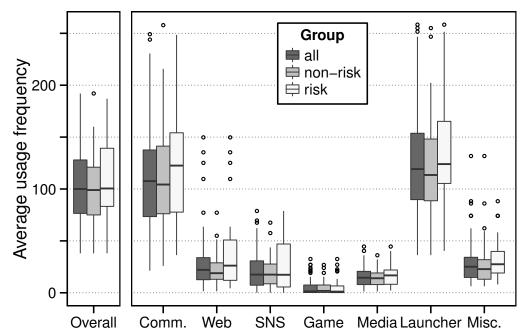

Figure 4. Usage frequency: overall and app-specific results

inspected the number of apps used during each session. We assumed that a session comprised a sequence of n apps used, $\{a_1, a_2, \cdots, a_n\}$ where a_k denotes the name of the kth used app, and the sequence length was simply given as n. Our results indicated that the sequence length was slightly longer for the risk group, but the difference was not significant (risk: 3.53 *vs.* non-risk: 3.16, p = .072, Cohen's d = 0.43).

We examined the app usage patterns of the risk group in order to determine whether the patterns were skewed. We checked the number of unique apps used during the experiment, but there was no significant difference (risk: 66.1 *vs.* non-risk: 65.5, p = .885, Cohen's d = 0.03). We used the entropy metric to measure the degree of usage for the top-k apps. For a given user, the entropy value was calculated based on the usage time distribution for the top-k apps using $-\sum_{i=1,k} p(i) \log p(i)$, where $p(i)$ is the relative usage time of the ith app. Entropy has the following property. The lower the entropy, the higher the level of focus on certain apps. For example, if a person only uses a single app, the entropy becomes zero. If she spends an equal amount of time on every app, the entropy is maximized.

Given that most participants frequently used a small number of apps, we examined the number of top-k apps using k = 5, 10, and 50, and we then calculated the entropy values. There was a significant difference in the top-5 app usage (p = .046, Cohen's d = 0.42), which showed that the top-5 usage patterns of the risk group were highly skewed. We performed unpaired t-tests (two-tailed) of the usage levels for the k-th ranked apps and found that the risk group spent more time interacting with the first ranked apps and on the second ranked apps (primarily KakaoTalk, Facebook, and browsers). The mean usage times of the first-ranked apps were 97.8 min and 69.9 min (p = .003,

Agg. Usage	Non-Risk [95% CI]	Risk [95% CI]	T	P	d
Usage time (m)	207.4 [187.3, 227.5]	253.0 [222.3, 283.8]	2.61	.011	0.54
Usage freq	100.1 [91.5, 108.6]	111.5 [97.2, 125.9]	1.47	.146	0.31

Session Usage	Non-Risk [95% CI]	Risk [95% CI]	T	P	d
Session time (s)	129.9 [115.6, 144.1]	157.6 [123.6, 191.6]	1.53	.134	0.36
Inter-ses. time (s)	816.6 [726.9, 906.4]	729.1 [623.7, 835.5]	1.25	.216	0.26
# unique apps	65.5 [60.0, 70.1]	66.1 [60.1, 72.1]	0.15	.885	0.03
Sequence len.	3.16 [2.9, 3.3]	3.53 [3.2,3.9]	1.84	.072	0.42
Agg. seq. len	204.2 [181.7, 226.6]	235.2 [204.4, 265.9]	1.66	.099	0.35
Top-5 entropy	1.96 [1.9, 2.0]	1.85 [1.7,1.9]	2.02	.046	0.42
Top-10 entropy	2.53 [2.5, 2.6]	2.40 [2.3, 2.5]	1.74	.085	0.36
#1 app time (m)	69.9 [60.6,79.0]	97.8 [81.5, 114.2]	3.02	.003	0.66
#1 app freq	115.4 [105.1,125.5]	134.7 [119.0,150.6]	2.18	.031	0.46
#2 app time (m)	37.2 [32.5, 42.6]	47.4 [38.4, 56.4]	1.93	.060	0.43
#2 app freq	70.3 [62.4, 78.1]	85.3 [70.2, 100.4]	1.78	.080	0.41

Table 2. Two-sample t-test results of the aggregated usage (two-tailed)

Cohen's d = 0.66), and the mean usage times for the second-ranked apps were 47.4 min and 37.5 min (p = .058, Cohen's d = 0.43), for the risk and non-risk groups, respectively.

Diurnal Usage: We hypothesized that the risk group would have different diurnal patterns compared with the non-risk group because they used smartphones for longer periods of time. We divided each day into four blocks; i.e., night: [0, 6), morning: [6, 12), afternoon: [12, 18), and evening [18, 24). We performed unpaired t-tests (one-tailed) of the usage levels with Bonferroni correction. Figure 5 shows that the risk group used their smartphones for significantly more time in the morning (p = .022, Cohen's d = 0.53) and evening (p = .038, Cohen's d = 0.49). Comparisons of the mean session durations across time blocks revealed that our participants had longer usage sessions during the morning and night blocks; e.g., morning *vs.* afternoon (p = .004, Cohen's d = 0.43) (detailed information can be found in the Appendix).

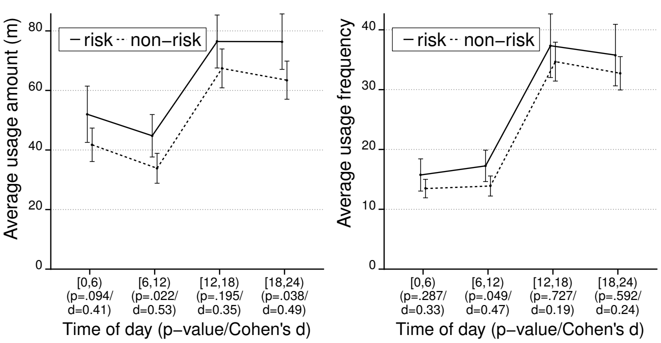

Figure 5. Overall diurnal usage time and frequency (mean value with 95% confidence interval, p-value/Cohen's d)

Summary of Findings: Our results showed that the risk group used their smartphones for more time than the non-risk group (risk: 253.0 min *vs.* non-risk: 207.0 min). The risk group users tended to use smartphones more frequently and to engage in longer usage sessions. Their app usage was highly skewed toward a small number of frequently used apps. In particular, significant differences were found in the usage amount/frequency of the first ranked apps. In addition, diurnal usage differences were observed, where the risk group used their smartphones for more time during the morning/evening than the non-risk group. Despite these usage differences, there was no significant difference in the number of unique apps used (risk: 66.1 *vs.* non-risk: 65.5).

Comm. Usage	Non-Risk [95% CI]	Risk [95% CI]	T	P	d
Usage time (m)	87.1 [75.8, 98.3]	98.8 [80.0, 118.6]	1.15	.257	0.24
Usage freq	112.5 [100.7, 124.2]	126.3 [107.4, 145.2]	1.32	.189	0.28
Voice time (m)	12.0 [9.4, 14.6]	14.5 [8.9, 20.0]	0.83	.411	0.20
Voice freq	6.2 [5.4, 7.0]	5.8 [4.6, 7.0]	0.51	.610	0.11
SMS time (m)	2.7 [2.1, 3.3]	4.4 [2.4, 6.4]	1.64	.108	0.41
SMS freq	6.9 [5.8, 8.0]	9.5 [6.7, 12.2]	1.75	.086	0.42
MIM time (m)	65.8 [56.1, 75.6]	75.6 [58.8, 92.3]	1.08	.281	0.23
MIM freq	76.9 [66.6, 87.2]	91.2 [73.9, 108.5]	1.53	.130	0.32
Inter-MIM time (m)	25.6 [19.8, 31.4]	21.0 [16.2, 25.8]	1.21	.228	0.23
Inter-MIM noti (m)	9.5 [4.9, 14.0]	6.9 [3.8, 10.0]	0.94	.351	0.17
MIM noti freq	378.5 [227.1, 529.8]	451.8 [449.1, 454.7]	0.64	.353	0.16

Web Usage	Non-Risk [95% CI]	Risk [95% CI]	T	P	d
Usage time (m)	41.1 [33.6, 48.6]	67.1 [48.4, 85.9]	2.60	.012	0.61
Usage freq	22.3 [18.7, 26.0]	38.5 [25.7, 51.3]	2.47	.018	0.61
Inter-web time (m)	81.0 [68.4, 93.5]	71.4 [53.3, 89.4]	0.90	.370	0.19

Table 3. Two-sample t-test results of category-specific usage (two-tailed)

CATEGORY-SPECIFIC USAGE PATTERNS

Our results revealed differences in aggregated usage between the risk and non-risk groups. In this section, we investigate these usage differences in detail by considering two of the most popular app categories: communications and web browsing. The overall results are presented in Table 3.

Communication App Use

Mobile Instant Messaging Usage: Figure 6 shows the usage times for different communication channels. Mobile instant messaging (MIM) dominated the overall communication usage, followed by voice calls, SMS, and emails. Because of the dominance of MIM usage, our analysis was focused on KakaoTalk, which is the most widely used MIM service in Korea. All of our participants, except one, used KakaoTalk on a daily basis.

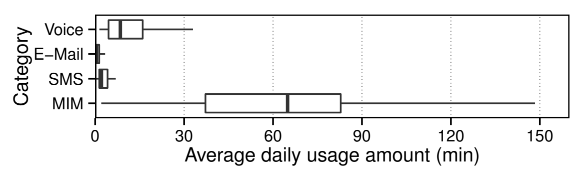

Figure 6. Usage statistics for communication apps

We calculated the mean daily usage time and frequency for KakaoTalk. The results showed that the risk group tended to use KakaoTalk slightly longer (risk: 75.6 min *vs.* non-risk: 65.8 min) and more frequently (risk: 91.2 *vs.* non-risk: 76.9), but significant differences were not found (see Table 3). To examine the KakaoTalk usage process more closely, we calculated the inter-app time (i.e., the time between consecutive KakaoTalk uses) and compared the mean inter-app time for the groups. The risk group had a slightly shorter mean inter-app time, but there was no significant difference (risk: 21.0 min *vs.* non-risk: 25.6 min; p = .228, Cohen's d = 0.23). We also calculated the inter-notification time (i.e., the time interval between two consecutive new notification arrivals). In Android, if a new message arrives when KakaoTalk is not an active app (hidden in the background or off-screen), an arrival event is placed into the notification drawer and a vibration or sound alert is generated. The inter-notification time was short (median = 26.6 s), but there was no significant difference (risk: 6.87 min *vs.* non-risk: 9.46 min; p = .351, Cohen's d = 0.17). Additionally, the number of notifications

		Externally Cued					Internally Cued					Total
		Kakao	FB	Email	Voice	Total	Kakao	FB	Email	Voice	Total	
Total num. sessions per day	Risk	38.9	5.4	0.5	6.0	76.7	*11.2	2.9	0.1	0.7	25.1	101.3
	Non-Risk	32.0	4.7	0.5	5.1	68.5	*7.2	1.7	0.1	1.2	18.8	89.0
Total app seq. length per day	Risk	*88.7	16.6	1.8	11.5	*199.6	14.9	4.5	0.1	1.3	36.2	235.2
	Non-Risk	*63.9	13.3	1.5	10.2	*164.6	11.2	2.6	0.2	2.4	31.6	204.2
Total session duration per day (s)	Risk	*4978.9	1084.7	112.0	305.8	*11910.7	540.5	272.0	1.4	14.3	1585.8	*13525.8
	Non-Risk	*3661.5	1010.0	112.3	251.8	*9645.8	385.4	134.2	10.2	36.6	1179.4	*11006.8
Per session duration (s)	Risk	170.1	233.9	209.6	58.8	390.4	55.3	69.5	3.0	31.0	64.3	335.5
	Non-Risk	164.6	263.2	170.8	57.6	348.4	64.7	70.1	30.7	24.9	70.6	314.2
App seq. length per session	Risk	2.6	3.2	2.9	2.1		1.4	1.4	0.2	2.0		
	Non-Risk	2.6	3.1	3.8	2.6		1.8	1.3	0.9	2.3		

Table 4. Usage differences in external and internal sessions (* $p < .05$)

per day was quite high, but the difference was not significant (risk: 451.8 *vs.* non-risk: 378.5; p = .353, Cohen's d = 0.16).

Notifications as External Cues for Usage: Smartphone usage can be triggered by external cues (e.g., incoming calls and messages), or internal cues (e.g., outbound calls/messages and web searches). In studies of technological addictions, external cues are regarded as a potential trigger of problematic usage behavior [18, 17, 10]. Given that our participants received a large number of notification alerts (more than 400 notifications; 90% of them from KakaoTalk), we investigated whether there were any usage differences among the externally cued sessions. We defined a session as externally-cued (or simply an external session) if any notifications from the apps that were used in the session have arrived during the time interval between that session and the preceding session (see Figure 7). Otherwise, sessions were treated as internal sessions.

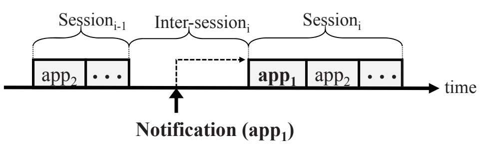

Figure 7. Illustration of an external session

Our data showed that the majority of sessions were external (79%). We further divided the external sessions based on the first app that was used in the respective session. For example, if an external session began with KakaoTalk, we referred to it as a KakaoTalk cued session. We considered KakaoTalk, Facebook, Email, and Voice Call as the first apps in external sessions because most notifications came from these apps.

For the external sessions, we found significant differences in the mean usage time per day (p = .037, Cohen's d = 0.44) and in the aggregated sequence length of the usage sessions per day (p = .033, Cohen's d = 0.45) (see Table 4). However, the number of sessions did not differ significantly (p = .192, Cohen's d = 0.28). We found no significant differences in these features of internal sessions. We examined the external sessions based on the first app (i.e., into KakaoTalk, Facebook, and Voice call cued sessions), and found significant usage differences only for KakaoTalk cued sessions with respect to the mean usage time per day (p = .030, Cohen's d = 0.50) and the aggregated sequence length of usage sessions per day (p = .029, Cohen's d = 0.50).

Summary of Findings: Our participants mainly used MIMs for mobile communications, and the usage was over 60 minutes per day. Each participant typically received more than 400 notifications per day on average and 90% of these notifications were from MIMs. There were no significant differences in the overall MIM usage, but MIM use tended to be longer (with the effect size of > 0.2) for the risk group and more frequent than the non-risk group, which was also supported by our diurnal usage analysis results (see Appendix). After dividing smartphone usage into internally-cued and externally-cued sessions, we found that 79% of the usage was external usage (e.g., MIM and Facebook). We found that the usage time of MIM-initiated sessions was significantly greater for the risk group compared with the non-risk group. This result confirms that MIM notifications act as external cues for smartphone usage and they can be considered to be a cause of problematic smartphone usage.

Web Browsing App Use

Usage Pattern Analysis: There were significant differences in the usage of web browsing apps between the groups (see Table 3). The daily usage times for the risk and non-risk groups were 67.14 min (SD: 55.25) and 41.14 min (SD: 28.87), respectively. Similarly, the daily usage frequencies for the risk and non-risk groups were 38.50 (SD: 37.77) and 22.30 (SD: 13.96), respectively. In addition, we compared the inter-app times of web browsers for the two groups. The risk group showed a shorter mean inter-app time: risk: 71.4 min (SD: 53.3) *vs.* non-risk: 80.9 min, (SD: 48.2), but the difference was not significant.

Content Consumption Pattern Analysis: We also compared the content consumption patterns of the risk and non-risk groups. We only considered the participants who used the default web browser, which allowed us to record their URL histories. As a result, there were 24 participants from the non-risk group, and 18 from the risk group. For each participant, we extracted the top ten frequently visited web sites and calculated their average usage frequencies per day. We aggregated these statistics within each group, and the results are presented in Table 5. We classified each domain name manually using the following categories: online communities, web portals (e.g., Naver and Yahoo!), news, web searches, entertainment (e.g., webtoons and movies), school web sites, and miscellaneous sites. Unequal-variance t-tests (one-tailed) showed significant differences in web portal usage (p = .049, Cohen's d = 0.60) and web search usage (p = .013, Cohen's d = 0.75). Web portals, such as Naver, Daum and Nate, are the top three web portals in Korea as a whole.

	Community	Portal	News	Search	Entertainment	School	Misc.	Total
Risk (SD)	43.7(101.7)	8.9 (14.8)	19.0 (59.2)	9.6 (9.2)	0.6 (0.8)	2.4 (3.8)	14.3 (11.3)	99.5 (130.6)
95% CI	[9.3, 78.1]	[3.9, 13.9]	[1.0, 39.1]	[6.5, 12.7]	[0.3, 0.8]	[1.1, 3.7]	[10.5, 18.1]	[55.3, 143.7]
Non-Risk (SD)	7.5 (13.2)	2.7 (3.1)	1.0 (1.4)	4.0 (5.3)	1.6 (3.6)	1.5 (2.1)	7.6 (11.6)	24.8(28.6)
95% CI	[4.0, 10.9]	[1.9, 3.5]	[0.6, 1.3]	[2.6, 5.3]	[0.4, 2.8]	[1.0, 2.1]	[4.6, 10.6]	[17.3, 32.2]
P (d)	.076 (0.53)	.049 (0.60)	.107 (0.46)	.013 (0.75)	.125 (0.42)	.189 (0.30)	.034 (0.57)	.014 (0.79)

Table 5. Distribution of the visit frequency on each category by the risk and non-risk groups

These web portals have similar content and functionalities, such as web search, news, trending issues, and links to other portal services. The risk group tended to check the portal pages for information updates more often (8.9 *vs.* 2.7 visits per day on average), and to search for information needs more frequently (9.6 *vs.* 3.9 visits per day on average) compared with the non-risk group. Detailed usage analysis of the Naver search queries showed that the risk group participants searched for trending issues (i.e., top search keywords) more often than the non-risk group participants.

We found that five participants who visited these sites to an excessive degree all belonged to the risk group. Three of these visited specific community sites 414.0, 166.3, and 96.7 times per day on average, respectively. Two other participants in the risk group visited news pages an average of 250.8 and 55.8 times per day, respectively. The most popular online communities among our participants included Todayhumor (sharing jokes, as on 9gag), Ruliweb (sharing game and animation information), and Ppomppu (shopping and discount information). These sites were visited primarily for the purpose of sharing user-generated content. We uncovered a unique usage pattern based on a detailed analysis of the URLs visited by our participants. Community web sites typically have a number of topical boards, and the participants would visit each separate board to check for new posts. Existing sites display new posts on a daily basis, but heavy web users checked these sites very frequently throughout the day, thereby exhibiting this pattern of checking for new information.

Summary of Findings: Compared with the non-risk group, our results showed that the risk group browsed the web more often and they tended to search for content updates more frequently (e.g., visiting web portals, and checking trending issues). Moreover, a few of the risk group participants searched for and consumed online content in an excessive manner and they exhibited unique surfing patterns while searching for this content.

ANALYTIC MODELING OF USAGE BEHAVIOR

Our results revealed quantitative differences in the usage patterns of the risk and non-risk groups. However, these results did not indicate how the usage features were associated with the key factors of smartphone overuse and how the usage features can be used to predict the potential for smartphone overuse. Thus, we perform analytic modeling of the usage behavior using multiple regression and machine learning.

Regression Analysis

The key factors (or symptoms) in the Smartphone Addiction Proneness Scale include interferences with work/personal activities (interference), virtual world orientation (virtual), emotional alterations such as anxiety and irritation if smartphone use is impeded (withdrawal), and habituated usage without reducing the level of usage (tolerance) [15]. Because of our limited sample size for regression analysis, we selected a set of representative features that characterized the general usage behaviors, including overall usage features (usage time/frequency, top-1 app usage time/frequency, top-5 app entropy, and usage time in 6-12/18-24 blocks) and category-specific usage features (MIM usage time/frequency, Web usage time/frequency, MIM/Web usage time in 6-12/18-24 blocks, external MIM session frequency, and aggregated MIM external session length/time).

	Total	Inter.	Virtu.	Withd.	Toler.
Model Summary					
Adjusted R^2	$.12^{**}$	$.14^{**}$	$.07^{*}$	$.07^{*}$	0.14^{**}
F	6.25	5.05	6.83	6.84	7.76
Standardized β					
Usage freq	$.26^{**}$			$.26^{**}$	
Top 1 app time		$.30^{**}$			
6-12 usage time			$.26^{**}$		
Web usage time	$.21^{**}$				$.31^{**}$
# ext. sessions					$.23^{**}$
Ext. MIM agg. seq. len.		$.48^{**}$			

$^{*}p<0.05$, $^{**}p<0.01$

Table 6. Regression analysis results

Table 6 summarizes the regression results. Overall, the usage time and frequency were closely related with smartphone overuse. Given that the usage of instant messaging was dominant, we hypothesized that its usage may play a critical role in problematic usage behaviors. Incoming MIM messages acted as external usage cues for smartphone use. The participants who experienced more interference tended to have longer session sequence lengths of MIM initiated sessions. Moreover, web usage and external cues were related to the tolerance factor; i.e., consuming online content that provides gratifications causes people to continue to use smartphones. These observations agree with the well-known theories of technological addictions [11, 17, 10]. Repeated content consumption (message exchanges, and web content) forms a habitual usage pattern, which may lead to addictive behaviors. External cues further reinforce this behavior.

Classification Analysis

We used machine learning algorithms to test the feasibility of the risk-group classification. During model training, we considered all of the features explored in the previous sections. For each feature, we considered the following basic statistics: mean, median, SD, and soft min/max (mean±2SD). We used the classification models in Weka v3.6 such as Decision Tree (DT), Naive Bayes (NB), and Support Vector Machine (SVM). We reduced the number of features using the Information Gain filter and Rank-based search algorithms. To determine the importance of the feature categories, we tested the performance using different category combinations (see Table 7). The general set contained the general usage features, such as the usage time/frequency, top-k entropies, and

Feature set	Acc. (%)	Pre.	Rec.	F-score	Classifier
All	81.05	.816	.811	.813	DT
General	72.63	.723	.726	.724	DT
Category	87.37	.874	.874	.874	DT
Temporal	78.95	.792	.789	.790	DT
ExtCue	64.21	.622	.642	.632	NB
ExcludeGeneral	85.26	.863	.853	.858	DT
ExcludeCategory	80.00	.806	.800	.803	NB
ExcludeTemporal	77.89	.782	.779	.780	DT
ExcludeExtCue	81.05	.816	.811	.813	DT

Table 7. Classification Results

sequence length. The temporal set included the diurnal usage features for aggregated usage and category-specific usage. The category set included the category-level usage features (no temporal aspects). The external set included all of the usage features for the external sessions. The models were evaluated based on a 10-fold cross validation. The dataset was divided into 10 roughly equal subsets. One subset was left out for validation, and the remaining subsets were used for training. The average performance of the classifiers were estimated by repeating this process for each subset.

Table 7 shows the best accuracy, precision, recall, and F-score values. We obtained the best performance when the category set was used with DT (F-score = 0.87). The ranked features in the best model included the web usage time (mean, SD, median), SNS usage time (SD), SNS usage frequency (soft max, median), inter-MIM time (median), and inter-browser time (mean, SD, soft max). By contrast, the performance was lower when all of the features were employed, where the ranked features included four general features, four category features, and two diurnal features. In addition, we examined the predictive power of each feature set. The performance of the general and temporal feature sets did not differ substantially, although the external feature set had the poorest performance. To measure the performance gain for each feature set, we trained the classifiers by excluding each set. As shown in the table, removing the category feature set resulted in a significant performance drop, which demonstrated its importance for classification. Moreover, the general features had negative effects on the overall performance.

In summary, we found that investigating various category-specific usage patterns was of critical importance, and our classification model allowed us to accurately classify whether a person belonged to the risk group. The current study focused mainly on communications and web browsing, but our feature selection results indicated the importance of other features. Thus, other categories such as social networking and mobile games may be explored in our future research.

PROBLEMATIC USAGE BEHAVIOR

We perform a content analysis of the interview/survey data to corroborate the earlier findings and to better characterize problematic usage behaviors.

Overall Usage Behavior

Our participants used smartphones for various purposes, which ranged from managing personal information and building social relationships, to passing the time and managing moods. Here, the seven participants who participated in the interviews are referred to using R (for the risk group) and N (for the non-risk group) followed by a number. Participants commented: *"In general, I use it to check news and updates. I can easily satisfy all of my curiosities with smartphones."* (N1); *"When I have nothing to do like waiting for someone, or if I feel bored during a class, I check Facebook."* (R1). The temporal usage patterns were closely related to the typical school lifestyle. Students have more free time in the morning (fewer classes), late at night (going to sleep), and over the weekends (no classes). During the day, smartphone usage was frequent and brief, as explained by N3: *"I frequently check my smartphone if there are new messages or alarms. I check it in every 5 to 20 minutes. I don't use it for a long time unless I play a game."* In general, our participants concurred that smartphone usage tended to last longer during the night, in the morning, or at the weekend (e.g., for relaxing and checking). For example, N1 stated: *"Unless I'm really tired, I always check my smartphone before going to bed. When I get up, I check my smartphone. Lots of updates [like Facebook] happen overnight."*

Frequent Interferences

We asked the participants an open-ended question to determine whether instant messaging interfered with their daily lives (in an exit-survey after the data collection process). The data showed that 92% experienced interference in various situations. One participant complained about a loss of attention and stated that: *"I sometimes lose track of a lecture due to KakaoTalk."* Similarly, another participant mentioned, *"I have to focus, but I check KakaoTalk in almost every 5 minutes."* Sleep patterns were disturbed, and one participant reflected, *"I sometimes chat before going to bed, and it makes me to stay up until late. Also, I was woken up by silly messages from my friends, who asked for game time via KakaoTalk."* Social activities were also interrupted often. One participant commented, *"When I'm dating or hanging out with friends, KakaoTalk messages make me feel nervous. I have had experiences where KakaoTalk disrupted ongoing conversations."* As demonstrated earlier, after receiving external cues, the risk group spent more time for using their smartphones. External cues disrupted the users' attention, and they may have experienced greater attention loss as their session durations became longer. Thus, the degree of interferences attributable to instant messaging was probably greater for the risk group than the non-risk group.

Habitual Usage and Limited Self-Control

The risk group expressed difficulty in regulating their smartphone use. They felt more compelled to check their smartphone: *"I keep paying attention, because I feel like new messages may have arrived."* (R1). Moreover, the risk group participants were less conscious of their smartphone use: *"I don't have any thoughts when using my smartphone. ... At that moment, [I'm] without any sense of time."* (R1). R2 reported an experience of excessive usage: *"It's not like I plan to use my smartphone, but I just turn on my smartphone unconsciously. I once used my smartphone to wake myself up in the morning. I got up at 9AM, but it turned out it was already 11AM."* R3 even claimed that smartphone usage is not problematic at all, by stating: *"I'm not a person who is* *likely to be addicted to something. I sometimes feel that I use my smartphone too much, but only very occasionally."*

The content consumption behavior of the risk group was less structured than that of the non-risk group, particularly when online content was consumed. In general, the risk group participants had difficulties in explaining the details of their content consumption behavior. R1 mentioned: *"I check articles in online communities in the morning, and I keep checking whether there are any updates. ... If one community does not have any updates, I visit other communities to check for updates."* R4 commented about Facebook usage: *"For example if there are no updates in Facebook, I check out a new person or keep following pages to seek for new content."* This behavior sometimes disturbed their sleep patterns. R4 acknowledged: *"I use my smartphone before going to bed, but in the end spend too much time; about one to two hours. There is always something new because I can dig in deeply. I don't sleep well because of this. After about two hours, I sometimes fall asleep even though I didn't think of stopping."*

By contrast, usage behavior was self-regulated in the non-risk group although they were also heavy users: *"I use my smartphone for about 20 minutes before going to bed. I take care of messages piled up in KakaoTalk, and check Facebook, and webtoon updates at 11:30 PM. After checking that, I'm done!"* (N1). N3 commented: *"I feel like Facebook is part of my daily task. I use Everytown [a game] when I feel drowsy after a meal, but not to fall asleep because it makes me feel lazy. I use Joara [an online novel app] when I have time to pass. I mostly use my smartphone, when I really don't feel like studying for an exam, when I have done all the work and there is nothing to do, and when I'm waiting for my friends, and I have some time to wait."*

DISCUSSION AND CONCLUSION

Our exploratory data analysis determined the usage patterns related to smartphone overuse. We found that the risk group spent longer time engaged in smartphone usage than the non-risk group. The number of apps used was similar, but the risk group exhibited a highly skewed usage pattern with respect to a few frequently used apps. In addition, there were significant diurnal usage differences, where the risk group used smartphones for longer periods in the morning and evening. Overall, our participants mainly used their smartphones for communications. MIM was the most frequently used app, followed by voice calls, SMS, and emails. They received a large number of MIM notifications per day (> 400). We found that the risk group spent more time on MIM-triggered sessions. This result serves as evidence that MIM notifications act as external cues that can lead to excessive smartphone usage. Another major component of smartphone usage was consuming various types of online content that provide instant gratifications (e.g., pastimes, entertainment, and information seeking). The risk group users spent more time on the web consuming these types of online content.

According to our analysis, the overall difference in the usage times between the risk/non-risk groups was not very high (< 50 minutes). In previous studies of technological addictions, however, researchers reported that the excessive and problematic usage depends mainly on function rather than the usage amount [11, 17]. For example, repeated usage for mood adjustment purposes (e.g., relieving boredom, stress, or depression) may form habitual usage and lead to addictive behaviors [11, 17, 10]. Our regression results demonstrated that smartphone overuse is closely related to the content consumption function of smartphones. Our interview results also provided detailed evidence of addictive usage behaviors. The risk group showed limited self-control, particularly when consuming online content (e.g., aimlessly following web/Facebook links while in bed). These findings provide new insights into previous research on scale development and problematic usage behaviors [16, 5, 13].

Understanding smartphone usage patterns has been an active area of research [8, 4, 25]. Our work supplements previous measurement studies [8, 4, 25] by reporting recent smartphone usage patterns (e.g., mobile instant messaging)—detailed information can be found in the Appendix—and by investigating problematic usage behaviors, which are emerging social issues associated with technology overuse. We observed significant changes in usage behaviors compared with the usage statistics based on the earlier datasets collected in 2009 [8] and 2010 [4]. We hypothesize that the much longer usage duration (224.7 min *vs.* 59.2 min [4]) may be attributable to the participant demographics, as well as the network effect of smartphones and MIM. Moreover, our research extends the study of Oulasvirta et al. [25] by examining real usage patterns from the perspective of smartphone overuse, and we demonstrated the importance of externally-cued usage behavior.

Our study provides new insights into the usage practices related to mobile communications [20, 3, 12]. We studied the latest trends for MIM use. As we reported, its usage is prevalent, and it can be indicative of overuse risks (e.g., causing negative impacts on an individual's personal and social activities). Thus, our results help to understand the impacts of semi-synchronous communication channels such as MIM, on the social expectations related to constant connectivity [3] and interruption management practices [20].

The use of computing technology for promoting health and sustainable behaviors has been of great interest to the HCI community [9]. In addition, addiction-related research has been increasing recently [26, 29, 30]. Our study on usage analysis and automatic behavior assessment may be useful when designing mobile software that can moderate excessive use (e.g., visualizing usage information, and utilizing social support), or it could facilitate the design of intelligent parental controls (e.g., addressing a child's smartphone overuse). We consider that designing and evaluating intervention software may be an interesting avenue for future HCI research [7]. Our research may provide a foundation for such research.

Similar to any cross-sectional and single-site study, the generalizability of this research may be limited by the characteristics of our participants. However, an earlier analysis partly affirmed the generalizability of this research [8]; i.e., fine-grained usage features such as the session time distribution exhibited consistent patterns across the datasets collected from different sites. We suggest that further exploratory and confirmatory studies might consider different sites, demographics, mobile devices/platforms, and cultural backgrounds. The present analysis could be extended by investigating additional contextual factors (e.g., location and activity), and content categories (e.g., social networking and mobile games).

ACKNOWLEDGEMENTS

This research was supported in part by the MSIP, Korea, and Microsoft Research, under the IT/SW Creative research program supervised by the NIPA (National IT Industry Promotion Agency) (NIPA-2013-ITAH0503130110270001000100100), by the IT R&D program of MSIP/KEIT [10041313, UX-oriented Mobile SW Platform], and by the National Research Foundation of Korea (NRF) grant funded by the Korea government (MSIP) (No. 2011-0018120, 2012R1A1A1008858). We thank the anonymous reviewers for their many useful comments and suggestions, which helped improve the paper.

REFERENCES

1. DSM-5. `http://www.dsm5.org/`.
2. The Mobile Consumer: A Global Snapshot, The Nielsen Company, Feb. 2013.
3. Ames, M. G. Managing Mobile Multitasking: The Culture of iPhones on Stanford Campus. In *CSCW* (2013).
4. Böhmer, M., Hecht, B., Schöning, J., Krüger, A., and Bauer, G. Falling Asleep with Angry Birds, Facebook and Kindle: A Large Scale Study on Mobile Application Usage. In *MobileHCI* (2011).
5. Carbonell, X., Chamarro, A., Griffiths, M., Oberst, U., Cladellas, R., and Talarn, A. Problematic Internet and Cell Phone Use in Spanish Teenagers and Young Students. *Anales de Psicologia* (2012).
6. Cui, Y., and Roto, V. How People Use the Web on Mobile Devices. In *WWW* (2008).
7. de Abreu, C. N., and Góes, D. S. *Psychotherapy for Internet Addiction*. Internet Addiction, John Wiley & Sons, Inc., 2011.
8. Falaki, H., Mahajan, R.and Kandula, S., Lymberopoulos, D., Govindan, R., and Estrin, D. Diversity in Smartphone Usage. In *MobiSys* (2010).
9. Fogg, B. J. *Persuasive Technology: Using Computers to Change What We Think and Do*. Morgan Kaufmann, 2002.
10. Greenfield, D. *The Addictive Properties of Internet Usage*. Internet Addiction, John Wiley & Sons, Inc., 2011.
11. Griffiths, M. D. Technological Addictions. *Clinical Psychology Forum* (1995).
12. Harmon, E., and Mazmanian, M. Stories of the Smartphone in Everyday Discourse: Conflict, Tension & Instability. In *CHI* (2013).
13. Hwang, K. H., Yoo, Y. S., and Cho, O. H. Smartphone Overuse and Upper Extremity Pain, Anxiety, Depression, and Interpersonal Relationships among College Students. *The Journal of the Korea Contents Association* (2012).
14. Kandell, J. J. Internet Addiction on Campus: The Vulnerability of College Students. *CyberPsychology & Behavior* (1998).
15. Kim, D., Jung, Y., Lee, J., Kim, M., Lee, Y., Kang, E., Keum, C., and Nam, J. Development of Smartphone Addiction Proneness Scale for Adults: Self-Report. *The Korean Journal of Counseling* (2012).
16. Koo, H. Y. Development of a Cell Phone Addiction Scale for Korean Adolescents. *J Korean Acad Nurs* (2009).
17. LaRose, R. *Use and Gratifications of Internet Addiction*. Internet Addiction, John Wiley & Sons, Inc., 2011.
18. LaRose, R., Lin, C. A., and Eastin, M. S. Unregulated Internet Usage: Addiction, Habit, or Deficient Self-Regulation? *Media Psychology* (2003).
19. Leung, L. Stressful Life Events, Motives for Internet Use, and Social Support among Digital Kids. *CyberPsychology & Behavior* (2007).
20. Levine, L. E., Waite, B. M., and Bowman, L. L. Mobile Media Use, Multitasking and Distractibility. *IJCBPL* (2012).
21. Miltenberger, R. G. *Behavior Modification: Principles and Procedures*. Cengage Learning, 2011.
22. Morahan-Martin, J. Internet Abuse: Addiction? Disorder? Symptom? Alternative Explanations? *Social Science Computer Review* (2005).
23. Morahan-Martin, J., and Schumacher, P. Loneliness and Social Uses of the Internet. *Computers in Human Behavior* (2003).
24. Murdock, K. K. Texting While Stressed: Implications for Students' Burnout, Sleep, and Well-Being. *Psychology of Popular Media Culture* (2013).
25. Oulasvirta, A., Rattenbury, T., Ma, L., and Raita, E. Habits Make Smartphone Use More Pervasive. *Personal Ubiquitous Computing* (2012).
26. Seay, A. F., and Kraut, R. E. Project Massive: Self-regulation and Problematic Use of Online Gaming. In *CHI* (2007).
27. Tossell, C. C., Kortum, P., Shepard, C. W., Rahmati, A., and Zhong, L. Getting Real: A Naturalistic Methodology for Using Smartphones to Collect Mediated Communications. *Advances in Human-Computer Interaction* (2012).
28. van den Eijnden, R. J., Meerkerk, G.-J., Vermulst, A. A., Spijkerman, R., and Engels, R. C. M. E. Online Communication, Compulsive Internet Use, and Psychosocial Well-Being Among Adolescents: A Longitudinal Study. *Developmental Psychology 44* (2008), 655–665.
29. Wang, K.-C., Hong, M.-T., Kao, C. H.-L., Lin, A. C., Wang, C.-A., You, C.-W., Pan, C.-H., Huang, M.-C., Liang, C.-M., and Chu, H.-H. A Phone-based Support System to Assist Alcohol Recovery. In *CHI EA* (2013).
30. Yarosh, S. Shifting Dynamics or Breaking Sacred Traditions?: the Role of Technology in Twelve-step Fellowships. In *CHI* (2013).

APPENDIX

We begin with summarizing the usage features that were used in this work. We then detail the descriptive statistics of various usage features that are important to understand the usage differences and similarities. First, we report the session duration and inter-session duration distributions, and show the diurnal patterns of usage sessions. We illustrate more detailed information about top-ranked apps (i.e., top-1, top-2) and report the distribution of the number of unique apps. Second, we plot the diurnal usage patterns of MIM and web browsing. Third, we report detailed information about internal and external sessions such as CI and *p*-values.

Usage feature summary: Table 8 summarizes all of the features used in this paper. The features can be divided into two groups: the overall usage and category-specific usage features. In the overall usage analysis, we investigated the aggregated usage (e.g., usage time, session frequency), session level usage (e.g., session time, inter-session time), and diurnal usage patterns. In the category-specific usage analysis, we mainly considered the communication and web browsing apps. Note that in our machine learning analysis, all of the categories were used; the features from the rest of categories are very similar to those of web browsing, i.e., usage time/frequency, and diurnal usage patterns.

	Sub-groups	Features
Overall Usage Patterns	Aggregated Usage	Usage time per day
		Session frequency per day
	Session-level Usage	Session time (usage time per session)
		Inter session time
		Number of apps used per session
		Entropy of top-*k* apps' usage time/frequency distributions (k = 5, 10, 50)
		Usage time and frequency of top 1/2 apps
	Diurnal Usage	Usage time, session frequency, and session time: night (0~6), morning (6~12), afternoon (12~18), and evening (18~24)
Category-Specific Usage Patterns	Comm.	Usage time and app frequency of communication, usage time and frequency of {MIM, SMS, email, call}, inter-MIM time
		Diurnal usage: usage time, app frequency, and session time of communication apps: night (0~6), morning (6~12), afternoon (12~18), and evening (18~24)
		Internal *vs*. external sessions: number of sessions per day, app sequence length per day, session duration per day/ session, app sequence length per session
	Web Browser	Usage time and app frequency of web, web content, inter-web time
		Diurnal usage: usage time, app frequency, and session time of web-browsing apps: night (0~6), morning (6~12), afternoon (12~18), and evening (18~24)

Table 8. Summary of usage features

Session-level usage: We constructed the group-level (or "average") distributions on session durations (Figure 8, 9) and inter-session durations (Figure 10, 11) by averaging users' sample percentiles, which is called Vincentizing, a nonparametric method. We plot the complementary cumulative distribution function (CCDF) and the cumulative distribution function (CDF). As shown in the earlier work [8], most sessions are short (median = 45.1 s). The frequency drops exponentially as the length increases, which is s by a near linear slope in the log-linear plot of CCDF in Figure 8.

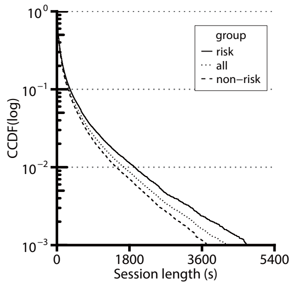

Figure 8. Session duration distribution (CCDF)

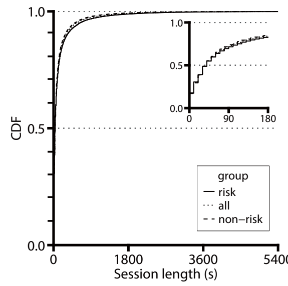

Figure 9. Session duration distribution (CDF)

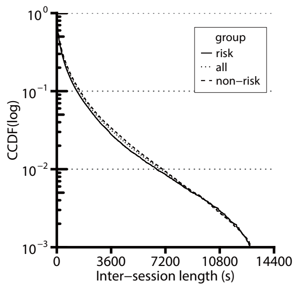

Figure 10. Inter-session duration (CCDF)

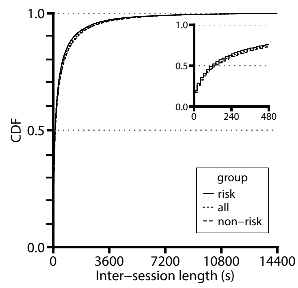

Figure 11. Inter-session duration (CDF)

The risk group tended to have longer session duration (risk: mean = 157.6 s, SD = 100.4 *vs*. non-risk: mean = 129.8 s, SD = 54.8) and shorter inter-session duration (risk: mean = 729.1 s, SD = 311.5 *vs*. non-risk: mean = 816.6 s, SD = 344.5), but we did not find significant differences (see Table 2). CCDF shows that usage of short sessions appears to be similar, but the risk group tended to have more number of longer sessions. We tested if the frequency of long sessions differed between the groups. Here, we simply assumed that a session was considered long if its duration was greater than 30 minutes. No differences were found in the aggregated usage duration of long sessions (risk: mean = 46.3 min, SD = 6.73 *vs*. non-risk: mean = 44.7 min, SD = 7.5, p = .293).

We calculate the mean session time during different time blocks, but there were no significant differences (see Figure 12). Most sessions are very short in both groups, and session durations are much longer during night and morning time blocks, when compared with those of the afternoon and evening time blocks. Table 9 summarizes the paired *t*-test results on the session durations between two time blocks.

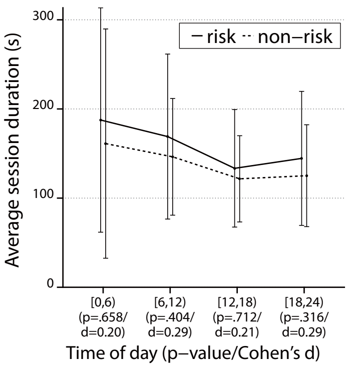

Figure 12. Diurnal usage session duration

Time Blocks	[0-6)	[6-12)	[12-18)	[18-24)
Mean [95% CI]	2.8 [2.4, 3.8]	2.6 [2.3, 3.3]	2.1 [1.9, 2.6]	2.2 [2.0, 2.8]
[0-6)		p=.289, d=0.15	p=.002, d=0.46	p=.009, d=0.38
[6-12)			p=.004, d=0.43	p=.031, d=0.31
[12-18)				p=.467, d=0.11

Table 9. Results of paired *t*-tests of the session durations between two time blocks

We plot the number of unique apps used by the participants in Figure 13. As shown earlier, no significant differences were found (risk: mean = 66.1 SD =17.7 *vs.* non-risk: mean = 65.5, SD = 21.0; p = .885, Cohen's d = 0.03).

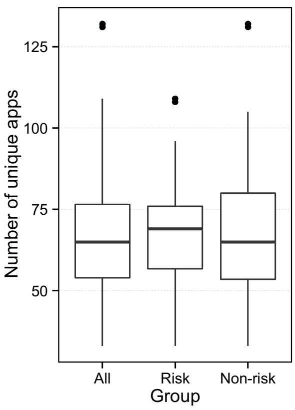

Figure 13. Average number of unique apps

Usage of top-ranked apps: We measured the degree of usage on top-*k* apps, by using the entropy measure. By varying the number of top-*k* apps as k = 5, 10, and 50, we plotted the entropy values in Figure 14. As shown earlier, a significant difference was found only in top-5 app usage (p = .045, Cohen's d = 0.42), showing that the risk group's usage patterns were highly skewed.

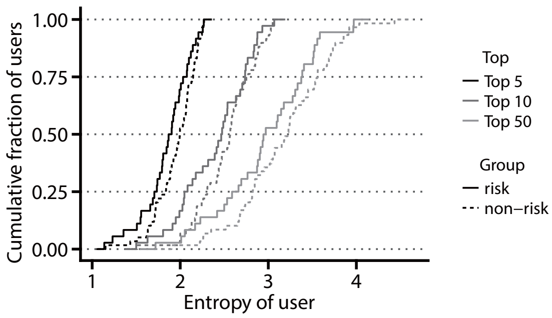

Figure 14. User entropy distribution (for top-*k* app usage distribution)

We also reported that the risk group spent more time on the top-*k* apps. Table 10 summarizes the category-specific distribution of top 1 and 2 apps. Top ranked apps belong to MIM, Web, and SNS; popular apps include KakaoTalk, Android's default web browser, and Facebook. In the case of other categories such as mobile games and multimedia, the apps are very diverse. The proportions of top-1/2 apps' category between the risk and non-risk groups did not show a significant difference (Top-1 apps: $X^2(5, N=95) = 4.750$, p = .447; Top-2 apps: $X^2(5, N=95) = 2.493$, p = .778).

	Top-1 Apps			Top-2 Apps		
	Overall	Risk	Non-Risk	Overall	Risk	Non-Risk
MIM	71.19%	58.33%	66.32%	27.12%	27.78%	27.37%
Web	20.34%	36.11%	15.79%	28.81%	22.22%	26.32%
Book	3.39%	0.00%	2.11%	0.00%	0.00%	0.00%
Game	1.69%	2.78%	2.11%	10.17%	8.33%	9.47%
SNS	1.69%	2.78%	2.11%	28.81%	38.89%	32.63%
Productivity	0.00%	0.00%	0.00%	1.69%	2.78%	2.11%
Multimedia	1.69%	0.00%	1.05%	3.39%	0.00%	2.11%

Table 10. Category-specific distribution of top-1/2 apps

Diurnal usage of MIM and Web apps: We plot the diurnal usage time and frequency of MIM and Web apps in Figure 15 and Figure 16, respectively. Similar to the overall usage patterns (as shown in Figure 5), there were significant differences in the morning (p = .017, Cohen's d = 0.64) and evening (p = .028, Cohen's d = 0.59). In the case of MIM, however, we did not find significant differences on the diurnal usage patterns. Given that the effect sizes of different time blocks are greater than 0.20, there are still possibilities of diurnal usage differences, and we suggest further studies on different datasets. One notable difference between MIM and web browsing was found on the usage time in the morning: when compared with MIM and web usage at night, the participants had much less time on MIM usage in the morning, whereas web usage was about the same or slightly increased.

Detailed information about cued sessions: In Table 4, we summarized the usage differences of internal and external sessions. In Table 11 and Table 12, we reported the detailed statistics, including CI, *p*-value, and Cohen's *d*.

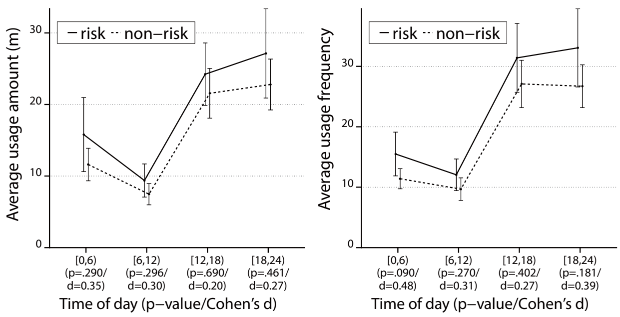

Figure 15. MIM diurnal usage time and frequency

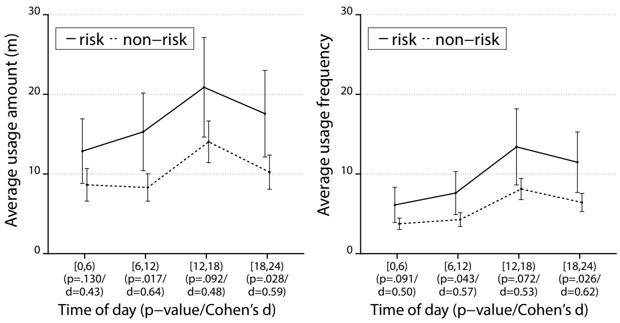

Figure 16. Web browsing diurnal usage time and frequency

		KakaoTalk		Facebook		Email		Voice		Total	
		Risk	Non-Risk	Risk	Non-Risk	Risk	Non-Risk	Risk	Non-Risk	Risk	Non-Risk
Number of session per day	Mean	38.9	32.0	5.4	4.7	0.5	0.5	6.0	5.1	76.7	68.5
	CI (95%)	30.1, 47.6	26.3, 37.6	4.0, 6.8	3.8, 5.6	0.2, 0.9	0.4, 0.7	4.4, 7.7	4.1, 6.0	65.4, 90.0	61.5, 75.5
	p/d-value	.170	0.29	.457	0.18	.924	0.04	.274	0.23	.192	0.28
Aggregated app seq. length per day	Mean	**88.74**	**63.90**	**16.6**	13.3	1.8	1.5	11.5	10.2	**199.6**	**164.6**
	CI (95%)	**69.1, 108.4**	**53.1, 74.8**	**12.0, 21.3**	10.6, 16.0	0.7, 2.8	1.0, 2.1	8.6, 14.3	8.2, 12.1	**145.9, 183.2**	**169.4, 229.7**
	p/d-value	**.031**	**0.50**	**.232**	0.28	.816	0.10	.428	0.17	**.038**	**0.44**
Aggregated session time per day (s)	Mean	**4978.8**	**3661.4**	**1084.7**	1010.0	112.0	112.3	305.8	251.8	**11910.7**	**9645.8**
	CI (95%)	**3944.8, 6012.9**	**3088.2, 4234.7**	**742.6, 1426.9**	801.2, 1218.8	41.6, 182.3	52.9, 171.8	193.3, 418.4	148.2, 355.5	**8399.6, 10891.9**	**10130.2, 13691.1**
	p/d-value	**.030**	**0.50**	**.721**	0.08	.997	0.00	.497	0.14	**.034**	**0.45**
Session duration (s)	Mean	170.1	164.6	233.9	263.2	209.6	170.8	58.8	57.6	390.4	348.4
	CI (95%)	130.1, 210.1	129.2, 200.0	175.9, 291.9	194.8, 331.5	131.2, 288.0	129.1, 212.6	43.3, 74.3	45.2, 69.9	301.3, 479.5	298.5, 398.3
	p/d-value	.842	0.04	.558	0.13	.602	0.22	.901	0.03	.409	0.19
App sequence length per session	Mean	2.6	2.6	3.2	3.1	2.9	3.8	2.1	2.6		
	CI (95%)	2.2, 2.9	2.3, 3.0	2.7, 3.7	2.7, 3.5	2.7, 3.1	2.5, 5.2	1.9, 2.4	2.2, 3.0		
	p/d-value	.759	0.07	.741	0.08	.342	0.21	.050	0.36		

Table 11. External sessions

		KakaoTalk		Facebook		Email		Voice		Total	
		Risk	Non-Risk	Risk	Non-Risk	Risk	Non-Risk	Risk	Non-Risk	Risk	Non-Risk
Number of session per day	Mean	**11.2**	**7.2**	2.9	1.7	0.1	0.1	0.7	1.2	25.1	18.8
	CI (95%)	**7.9, 14.6**	**5.2, 9.2**	1.5, 4.1	1.2, 2.3	0.0, 0.1	0.1, 0.1	0.5, 1.0	0.4, 1.9	17.0, 33.3	15.0, 22.6
	p/d-value	**.031**	**0.46**	.125	0.38	.720	0.16	.254	0.20	.160	0.34
Aggregated app seq. length per day	Mean	14.9	11.2	4.5	2.6	0.1	0.2	1.3	2.4	36.2	31.6
	CI (95%)	10.4, 19.4	7.4, 14.6	2.3, 6.3	1.8, 3.5	0.0, 0.2	0.1, 0.3	0.9, 1.7	1.0, 3.7	25.2, 47.3	24.3, 39.0
	p/d-value	.218	0.26	.122	0.38	.694	0.17	.134	0.26	.469	0.15
Aggregated session time per day (s)	Mean	540.5	385.4	272.0	134.2	1.4	10.2	14.3	36.6	1585.8	1179.4
	CI (95%)	345.1, 735.9	253.1, 517.7	110.5, 415.9	88.2, 180.2	0.1, 2.6	0.2, 20.6	10.0, 18.7	9.9, 63.3	991.1, 2180.6	917.1, 1441.6
	p/d-value	.177	0.29	.124	0.40	.259	0.25	.104	0.27	.211	0.30
Session duration (s)	Mean	55.3	64.7	69.5	70.1	3.0	30.7	31.0	24.9	64.3	70.6
	CI (95%)	39.2, 71.4	45.8, 83.5	47.2, 87.3	47.6, 92.6	0.03, 5.8	19.5, 41.9	16.5, 45.6	14.1, 35.8	47.7, 80.8	56.7, 84.6
	p/d-value	.451	0.15	.973	0.04	.003	0.69	.495	0.15	.562	0.12
App sequence length per session	Mean	1.4	1.8	1.4	1.3	0.2	0.9	2.0	2.3		
	CI (95%)	1.2, 1.6	1.3, 2.3	1.1, 1.7	1.6, 1.4	0.0, 0.4	0.7, 1.2	1.6, 2.4	1.7, 2.8		
	p/d-value	.127	0.26	.463	0.11	.114	0.67	.421	0.15		

Table 12. Internal sessions