Research Article

Getting Real: A Naturalistic Methodology for Using Smartphones to Collect Mediated Communications

Chad C. Tossell,[1] Philip Kortum,[1] Clayton W. Shepard,[2] Ahmad Rahmati,[2] and Lin Zhong[2]

[1] *Department of Psychology, Rice University, 6100 Main Street, MS-25, Houston, TX 77005, USA*
[2] *Department of Electrical and Computer Engineering, Rice University, 6100 Main Street, MS-25, Houston, TX 77005, USA*

Correspondence should be addressed to Chad C. Tossell, chad.tossell@rice.edu

This paper contributes an intentionally naturalistic methodology using smartphone logging technology to study communications in the wild. Smartphone logging can provide tremendous access to communications data from real environments. However, researchers must consider how it is employed to preserve naturalistic behaviors. Nine considerations are presented to this end. We also provide a description of a naturalistic logging approach that has been applied successfully to collecting mediated communications from iPhones. The methodology was designed to intentionally decrease reactivity and resulted in data that were more accurate than self-reports. Example analyses are also provided to show how data collected can be analyzed to establish empirical patterns and identify user differences. Smartphone logging technologies offer flexible capabilities to enhance access to real communications data, but methodologies employing these techniques must be designed appropriately to avoid provoking naturally occurring behaviors. Functionally, this methodology can be applied to establish empirical patterns and test specific hypotheses within the field of HCI research. Topically, this methodology can be applied to domains interested in understanding mediated communications such as mobile content and systems design, teamwork, and social networks.

1. Introduction

Smartphones have provided ubiquitous computing and communication resources to a growing number of users. The International Telecommunication Union [1] recently reported over 940 million smartphone service subscriptions worldwide and that this number is growing exponentially. These devices have transformed the mobile phone into a technological companion [2] that is completely portable, connected, available, and powerful. Researchers can now leverage logging technology available through these devices to access real communications data from real environments [3].

However, those employing this technology must guard against the implicit assumption that logging does not affect the behavior of participants [4]. Similar to traditional field methodologies, reactivity (i.e., a modification in behavior as a consequence of being measured [5]) can occur if careful steps are not taken to plan for the "selection, provocation, recording and encoding of behaviors and settings" [6] in a way that preserves naturally occurring behaviors [7]. This could seriously impact both internal and external validity of the data collected via these devices [7]. We submit, with careful design, smartphone logging can be used for enhanced naturalistic studies to better establish empirical patterns, develop theories, and test specific hypotheses in communications research [8]. This technique seems ripe for human factors (HFs) domains such as interface design, teamwork, and social networks to collect and analyze an enormous amount of mediated communications data from real settings.

The present work contributes a naturalistic use of smartphone logging to preserve realistic behaviors. To this end, we begin with a review of relevant literature to demonstrate some of the capabilities and applications of the emerging technique. Second, we describe the method in detail and include some of the important constraints that must be considered when implementing the methodology. Third, we describe some of the specific benefits and limitations of

the methodology by way of examples. We conclude with a discussion of how the methodology fits into HCI research in general.

2. Background

Smartphones are used in diverse settings [9] to share information within and across teams [10], maintain social relationships [11], and develop social networks [12] among a number of other things. To understand communications through these devices, traditional research methods (e.g., laboratory, field) are commonly used or adapted to fit mobile environments [13]. Each of these methods offers several benefits and limitations [14, 15]. For instance, laboratory studies are highly controlled and can provide data high in internal validity [16]. The potential drawback, however, in many laboratory studies is the lack of ecological validity due to the artificial setting [17].

When studying interactions with smartphones, traditional field methodologies (e.g., diary studies, ethnography, observation) seem to be a more natural fit to enhance ecological validity [18, 19]. However, field studies have met challenges when applied to studying communications outside of the lab; two are briefly described here. Observer effects are a primary concern. For example, Eagle and Pentland [20] described several examples where invasiveness adversely impacted the validity of communications data. These field studies required researchers to view confidential meetings, teenagers in their bedroom, and the communication of lovers. Second, traditional data collection techniques have largely required user inputs which may adversely influence accuracy. Diary studies, for example, often interrupt users from their main task, place a burden on participants to report, and rely on their memory of events [21].

Logging methodologies have addressed many of these concerns by allocating observations to technology. These methodologies provide access to data that can be collected without an observer present or a requirement for users to provide self-reports [3]. Tasks do not have to be constructed by the experimenter. Instead, data can be pulled from participants' daily activities on familiar interfaces within normal contexts [22]. Thus, data collected from loggers are typically considered more objective, accurate, and realistic [20, 23].

Using logs to understand realistic behaviors with technology is not a new approach [15, 24]. More recently, researchers in HCI have used Web logs to better understand browsing strategies [25], search behaviors [26], revisitation of websites [27], and usage differences between groups (e.g., novice-expert [28]). These studies have characterized interaction behaviors for enhanced design of interfaces [29]. For instance, website revisitation rates have been analyzed in multiple studies [25, 30] and applied to the design of history interfaces for internet browsers.

In research on teams, communications data have been collected and analyzed via automated methods to assess performance, develop training, and reduce errors [31]. Both physical communication data and the content of communications have been analyzed to understand team cognition [32] and performance [33]. Much of the work conducted in this domain has collected communications content relevant to team tasks on dedicated systems (e.g., the radio on an airplane [33]) and applied to training development and the design of communication systems. Smartphones, in contrast, are used across tasks for both personal and professional communications [10].

Smartphone logging of communications data is a recent trend. Most notably, authors of [20] installed logging capabilities on first-generation smartphones to passively collect data from participants in two academic departments. These data have been used to understand the development of social networks [12, 34, 35], predict smartphone usage [36], classify behavioral patterns [37], and study the role of location [20] and temporal patterns [38] in communication behaviors.

More intrusive and controlled studies have also been conducted using smartphone logging. Authors of [39] sent messages to participants at various times and asked participants to take pictures of their current setting and describe their information needs. These responses were logged and analyzed for technology design. Authors of [40] used a within-subjects experimental design to assess the effects of a novel interface on decreasing missed phone calls. This study used a longitudinal approach and collected a large amount of data via smartphones to convincingly show the effectiveness of the new system. Clearly, smartphone logging has been employed in a number of diverse ways.

Although logging has been applied effectively to a number of research aims, the technology has not been advantageously employed in a systematic manner to preserve naturalistic behaviors. For instance, many of the previous logging studies mentioned above have reminded participants they are being measured by requiring them to report data, introducing novel interfaces, or collecting data considered private. Similar to other research methodologies, using logger technology involves planned constraints to improve accuracy [6]. We submit that employing loggers can be more naturalistic by reducing the potential for provoking normal behaviors due to measurement. Removing this unwanted variance is clearly important for all research applications. Below we describe the constraints involved with implementing smartphone logging and how we employed a more naturalistic approach to study communications.

3. Collecting Naturalistic Communications

One of the central tenets in Human-Computer Interaction (HCI) is to understand user differences based on demographics, experience, or other characteristics [41]. To date, there is a lack of studies using naturalistic and longitudinal methodologies to assess these user differences for enhanced mobile systems design. Because smartphones are becoming more ubiquitous, we think such methodologies could be beneficial for the future design of mobile systems and content. This section considers a number of factors important to implement a naturalistic approach to logging smartphone usage. Decisions on each of these factors can influence the level of realism of data collected. Our methodology requires the researcher to address nine considerations in the design of the mobile-logging-based studies (Table 1). These

Table 1: Nine constraints for designing a study using logger technology.

Label	Consideration
Variables	What variables are needed? Logging is selective and intentional regarding what is collected and what is ignored. For instance, researchers can collect a large number of contextual and demographic variables to a small number (e.g., just time and type of searches). Common dependent variables include physical data alone (e.g., call frequency & duration) to physical and content data (e.g., SMS message).
Privacy	Are data required potentially sensitive to participants? Many actions performed on smartphones are considered private.
Obtrusiveness	How do I collect the data? This can range from fully automated (low interruptions) to requiring participants to report (e.g., experience sampling with logger).
Interface	What interface(s) will participants use? New interfaces can be introduced or logging can be embedded and run as a background process on current interfaces.
Tasks	What tasks will participants perform? These tasks can be completely naturalistic (i.e., participant constructed) or experimenters can construct artificial tasks.
Technology	What technology is used? Logs can be pulled from public files (e.g., search databases) which would allow participants to use familiar technology. On the other end of the spectrum, researchers can provide new instrumented technologies to participants.
Participants	Who are the participants? Subjects may consist of a random population of people that are totally unaware they are being studied to individuals within an academic department or domain of interest (e.g., pilots) that are highly aware of the measurement.
Setting	Where will the study take place? One benefit of smartphone logging is that communication data can be collected in real environments (instead of a laboratory).
Study duration	How long to measure usage? This could range from one task of interest to longitudinal measurements over a period of months or years.

considerations are not necessarily unique to smartphone logging. Some are shared by other methodologies, while others apply more directly to smartphones. We describe each consideration in turn and how decisions can impact reactivity. Of course, many of the factors are not mutually exclusive (e.g., privacy).

Variables. As with any research endeavor, the variables of interest highly influence the nature of the methodology. A naturalistic approach to logging must take several precautions in selecting data to be collected. Foremost, mediated communications, such as text messages, are considered more private than mail [42]. By collecting data such as these, participants could change their normal communications behaviors. Researchers implementing a naturalistic logging methodology can, however, record physical data from communications such as word count. Additionally, researchers that have a particular interest can collect a select amount of content. For instance, in one study, we hashed all communications data except emoticons. Table 2 is an example of some of the data that have been collected through naturalistic smartphone logging.

Privacy. Privacy must be considered at multiple levels in a naturalistic logging methodology. Smartphone loggers can collect the content of communications including from people outside of the study [43]. Collecting these data, however, can adversely impact user behavior because users may be reluctant to engage in highly personal communications if their privacy is not guaranteed [24], even though they may adapt [44]. Capturing communications content may not be considered invasive for dedicated professional systems [31]; however, communications data on smartphones are highly private in nature [42].

Several privacy constraints should be implemented throughout a study in order to conform to the methodology. First, we submit that participants should be aware of how their data are to be used. The rationale of the study and anonymization process should be explained to participants in detail before the study begins. This process at the beginning of the study has also been noted as an important step to minimize reactivity [44]. Second, participants should be assigned participant numbers to keep user interactions anonymous and researchers must be intentional about avoiding linking usage data with names. This can

TABLE 2: Example of data that can be collected through logging.

SMS and e-mail	Phone	Web	Apps
Contact	Contact	URL	Name
Date/time	Date/time	Date/time	Date/time
Duration	Duration	Duration	Duration
Launches	Outgoing or ingoing	Referring URL URLs	
Sent or received			
Word count			
Type of emoticons			
Number of Emoticons			

be particularly challenging when the study is longitudinal due to concerns such as phone malfunctions. For instance, in one implementation of the methodology, constraints were designed beforehand to maintain privacy for times when researcher-participant interactions were required (e.g., phone malfunctions). Only one researcher, who did not have access to data directly from the iPhones or the server, interacted with participants. Because this researcher did not have knowledge about their participant numbers, no data could be linked to the malfunctioning device. The technical issues were passed to other researchers along with the phones without any information that could identify the participant. Only phones were matched to user IDs in order to prevent any linkages between user IDs and names.

At another level, the logging technology can be designed to help preserve users' privacy. For example, participant numbers should be automatically associated with usage data on the phone and an encrypted tunnel should be used to transfer the collected data, in order to prevent unauthorized eavesdropping. The logger should not allow researchers or participants to view the actual content and contact information of emails, text messages, phone calls, and the address book. Instead, researchers can employ other methods to retain research critical data without collecting sensitive data. Contact information (i.e., phone numbers, names, and e-mail addresses) can be automatically assigned unique alphanumeric codes by the logger before it reaches any human. Similarly, text analysis, performed on the device, can extract relevant information from communication content (e.g., word count) and return only that information, not the specific content. By employing these measures, no potentially sensitive information ever leaves the phone, but important data can be linked together for analyses. For example, in a study where a participant sends a text message to her mother and then calls her mother later, the same code should be assigned to the contact for both transactions. Although content information could have been captured, this can negatively influence the realism of user behaviors. Steps such as these can enhance user privacy, and, subsequently, more naturalistic data can be collected because participants' normal behaviors are not disrupted because of privacy concerns.

Participants. A small number of HCI studies have reported data from users unaware that they were being recorded. Many times these include data from a cookies log that were recorded from websites, such as search engines. To implement the current methodology, participants should be fully informed of the data being collected from their phones. Though this may increase reactivity (at least initially), research ethics should also be taken into account and considered in the design of the methodology. Careful selection of experimental participants is also important. For instance, researchers should avoid recruiting subjects with previous relationships to the experimenters or the design of the study. It is also important to limit the potential reinforcement that the participants are being measured by minimizing or eliminating non-study-related meetings. Additionally, other participant behaviors, such as international travel, should be considered. If cost is not an issue in the administration of the study, then travel may not be an issue. If cost is an issue, then selecting participants who do not have extensive overseas travel planned would be important to maintaining the completeness of the data.

Study Duration. Longer studies allow the effects of being measured to wear off [45, 46]. Additionally, some events are by their very nature of low frequency. Small time frames might miss key events. Further, longer data collection efforts have the potential to yield richer information about cycles and trends that might not be obvious in shorter studies. Although there are no strict suggestions for the duration of smartphone logging in order to apply this naturalistic methodology, longer is generally considered better. We suggest that researchers also consider other factors such as familiarity with the device when choosing the study duration.

Obtrusiveness. Similarly, measurement obtrusiveness increases participant reactivity [5]. There are a number of ways researchers or logging technology can intrude and remind participants they are being watched or impede on normal behavior. For instance, requiring users to respond to text messages or perform a data upload procedure to collect data (e.g., performing an online action) can increase subject reactivity. While these kinds of activities can provide valuable information such as the immediate context where users are using their device and other self-report information, they come at the cost of interrupting normal activities. In addition to interrupting normal behaviors, constant requirements for users to perform any study-related actions

are both unnatural actions by nature and may residually lead to additional activities that might not normally occur. A naturalistic logging methodology should not require *any* user actions to record data.

Beyond the technical design of the logger, a minimum number of participant contact meetings should be scheduled with participants to collect self-report data once the study has commenced, if at all. An optimal implementation of the current methodology is for research-related meetings to be scheduled before logging begins and then after logging ends. Any meetings during usage data collection could again remind users they are being measured and this provocation could lead to reactivity.

Interface. Another factor important to consider in preserving realistic and generalizable behaviors is the types of interfaces implemented on the technology being used in the HCI study. For instance, employing novel interfaces (e.g., a custom browser) or changing technologies over the course of the study (e.g., phone swapping) can adversely affect data validity by producing false rates of behaviors, increased variability and result in numerous other problems [45]. Users habituate to being measured over time with a stable interface [46]. Constant reminders that the technology is being logged simply reinforce the feeling of being observed, much like a live observer can adversely impact subject behaviors [47].

As with most of these considerations, tradeoffs must be made. In web logging studies it has become common to require users to install and use a different browser with a unique interface, in order to capture interactions such as button clicks (e.g., the back arrow) and use of history systems (e.g., bookmarks). The down side, of course, is that the ability to generalize these results may be problematic because interactive behaviors may have been driven by the novel interface and not what the users would have normally done on their usual browsers.

Tasks. The tasks users perform can range from self-constructed tasks in ecologically valid environments to researcher-constructed tasks in controlled laboratory environments. Of course, there is value to each approach. The latter approach applied to smartphones can be used to achieve statistical control and assess specific HCI problems (e.g., usability for common tasks [48]). The external validity of such studies, though, may be questionable due to the highly contextual nature of smartphone use. A naturalistic approach allows users to perform the tasks they might usually do with their smartphone. To apply the current methodology, researchers should avoid influencing what users do with their smartphone.

Technology. One challenge in smartphone logging is the design of constraints to encourage participants to use the instrumented technology as if it was their own. This can be difficult because smartphones are not typically used in isolation of other technologies (i.e., "on an island" [49]). Many actions that can be performed on smartphones can also be performed on other technologies such as a laptop or another flip phone. For the current methodology, we suggest researchers require their participants to use the instrumented smartphone as their primary device and provide incentives to encourage this behavior. Thus, researchers should provide smartphones that represent the latest commercial offerings to promote this transition or work with phones previously purchased by the participants. Further incentives, such as unlimited data, texting and copious nationwide phone minutes, can further entice participants to use the experimental equipment exclusively.

4. Applications

Clearly, two of the primary benefits of naturalistic logging are the tremendous amount of data that can be collected and access to data not meant for the public eye. This section describes how these data differ from self-reports, provides evidence of preserving realistic behavior, and details several example applications of a naturalistic smartphone logging methodology applied to research problems in HCI.

4.1. Example Applications. Data sets obtained from logging smartphones can be extremely large. Indeed, in studies conducted by our lab using this methodology [50, 51] the amount of data gathered was enormous. For example, for a population of 24 participants over 18,000 hours of iPhone usage was captured. This included over 650,000 application launches, 460,000 sent and received text messages, and 42,000 phone calls. Although providing new smartphones along with free service may seem costly, the amount of data received in return is large.

A naturalistic smartphone logging methodology can be applied to a number of research problems. We briefly provide two examples of studies that leveraged the proposed methodology. In the first, we examined gender differences in emotive expressions online [50]. The second example is a snapshot of data collected over a period of one year to characterize communications through SMS and voice phone channels.

4.1.1. Emoticon Use. In [50], we had a particular interest in emoticon use through text messages (SMS). We obfuscated the content of the text messages and the contact information between users. However, we recorded both the number and type of emoticons sent and received by our participants. We used these data to examine differences between genders in their use of emoticons. Our naturalistic logging approach examined a smaller number of users over a period of six months. Still, reliable differences were found between genders. Contrary to previous studies that suggested technology closes the gender gap, our results showed that females more frequently used emoticons within text messages. The number of emoticons sent via participants' phones was adjusted by number of messages and verbosity. On all counts, females sent and received more emoticons. Surprisingly, however, emoticon vocabulary ratios calculated for each participant (number of unique emoticons sent/total number of emoticons sent) revealed that males sent out a wider range of emoticons compared to females.

TABLE 3: Percentages of time spent on iPhone applications from both self-reports and actual usage.

	Duration		Frequency	
	Self-report mean %	Actual usage mean %	Self-report mean %	Actual usage mean %
SMS	26%	14%	32%	27%
Phone	17%	16%	11%	14%
Email	19%	4%	21%	10%
Web	18%	6%	19%	6%
Other	20%	61%	17%	43%

Previous studies that analyzed emoticon differences were mixed. All of these data, however, were from public content (e.g., listservs, blogs, etc.). When private data were analyzed (after gaining full consent from participants), stable gender differences were found. These results could be applied to the design of future smartphone communications systems, for instance, providing easier ways to personalize smartphone keyboards to allow some users (e.g., females) to surface frequently used characters (e.g., happy face emoticons).

4.1.2. Characterizing Communications in Social Networks. In another study, we explored how text messaging (SMS) and voice phone mediums were employed to encounter contacts in participants' social networks. Over 42,000 phone calls and 346,000 text messages were collected between 5,291 participant-nonparticipant dyads that made up our dataset for this analysis. Zipf-like distributions were found within each of these modalities. Thus, a small number of contacts were encountered very frequently via text messaging and voice phone calls and a long tail of contacts were contacted once or twice.

To understand empirical patterns associated with how both communication modalities were used, we examined longitudinal patterns. 24% of the contacts were encountered by our participants via both modalities (these contacts are referred to as "intermodal contacts" hereafter). 57% of contacts were encountered on voice phone only. The remaining 19% were encountered on SMS only. We also observed high stability of contacts encountered across both modalities. 96% of intermodal contacts were encountered across at least two months. 71% were encountered over 7 months. Thus, many of these intermodal contacts were likely more strongly tied to our participants.

Findings in communication patterns with each of these contacts revealed optimization trends that cannot be captured in any other way. The number of messages sent to intermodal contacts increased over time. These messages were shorter in length compared to messages sent to other contacts. Conversely, the number of phone calls made to these same intermodal contacts decreased over time. The duration of these calls was generally higher than phone calls made to other contacts. Future interfaces could be designed to better integrate aspects of these two modes of communication. For instance, a more intelligent linkage between the two modes might show an integrated history of communications with contacts. Another potential improvement would be

for SMS applications to better transfer draft messages that require higher data entry as the interface on smartphones seems to afford shorter messages.

4.2. Assessments of the Methodology. Studies conducted using logging technology result in the enhanced accuracy of the resulting data compared to other methods [20]. This has been corroborated in our own studies using the methodology. Participants were asked to report the relative proportions of usage in terms of both frequency and duration for five categories (Table 3). We found that users were fairly accurate at ranking how much they used various communication applications. However, they were not precise at estimating the relative amounts. Participants significantly underreported the amount of time spent on other applications. This is a vivid demonstration that smartphone logging can be used to collect data that cannot be collected accurately via other means (e.g., self-reports [53]).

One of our studies that applied the above methodology has also indicated that continuous measurements do not influence participants' normal behaviors. Participants responded on a Likert Scale (1 = strongly disagree to 5 = strongly agree) to the following statement: "The fact that my iPhone use was being measured changed my normal behaviors." 89% of the participants strongly disagreed and no one responded with a 4 or 5 ($M = 1.17$). A similar open-ended question was also answered and 86% indicated no change in behavior, 8% indicated that measurements initially changed their behavior, but that the effect quickly faded, and 8% indicated that it affected specific behaviors such as application downloads.

5. Discussion

Logging is a flexible approach that has addressed many of the challenges associated with traditional research methods [3, 13, 20]. However, researchers must take great care in designing these kinds of studies to preserve realism and minimize the potential adverse effects of measurement on behavior. We introduced nine important dimensions to consider in this regard: privacy, variables, interface stability, technology selection, nature of the task(s), participant selection, setting, and duration of the study. Decisions on these elements contribute to the overall realism of behaviors captured. The overall level of realism is important because it allows researchers to collect data without provoking normal behaviors.
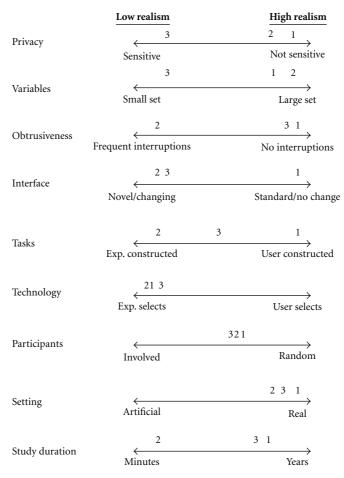
FIGURE 1: Three studies (1-current study, 2-experience sampling [52], 3-experimental [40]) that used smartphone logging to collect data and the approximate placements on each of the nine considerations.

Of course, many of these dimensions have trade-offs. Striking a balance between collecting relevant data and impacting real behaviors can be challenging (similar to other methodologies). The pursuit of specific research goals may mean that it is not always necessary to have the highest level of realism. For example, if the goal is to understand the role of location on smartphone usage that uses semantic analysis of messages [43], then privacy constraints could be relaxed in order to accomplish this goal. And if novel interfaces are the subject of the research [40], then the importance of the stability of the interface would need to be relaxed to assess the effects of the interface. Logging can also be employed in an obtrusive way (e.g., experience sampling [39]) to get more qualitative information (e.g., pictures of current location).

The primary strengths of the methodology introduced in this report are the commitment to naturalistic data collection and longitudinal nature of the study. Figure 1 shows where three different studies (including the current one) roughly fall along the dimensions introduced above in Table 1. The first study is the current methodology described in detail in this paper. The second study was conducted by Jönsson et al. [52] and used SMS probes to collect information on learning environments for development of distributed pedagogical tools. These probes were sent everyday to students' mobile phones at an unpredictable time with instructions. These instructions were in the form of a game or request and had students use their provided phones to collect information (e.g., take a picture of your surroundings). This is similar to other studies that used smartphone logging for participatory design (e.g., [39]). The third study conducted by Oulasvirta et al. [40] used a repeated-measures approach combined with unobtrusive logging to understand communications via smartphones. A combination of experimental control using a standard A-B intervention, a longitudinal collection period (265 days), and the collection of a host of contextual and usage variables truly demonstrates the innovative methods that can be employed with logging [3]. They also recorded voice phone communications for qualitative analyses. Decisions on each consideration can vary widely across studies, confirming that smartphone logging is a flexible tool that can be leveraged in a number of ways based on research goals. The current approach rates high on most of the dimensions, as seen in Figure 1 suggesting that the behaviors measured were more realistic.

Of course, although logging in a highly realistic fashion can yield a wealth of information, there are limitations. For example, the technology cannot directly capture user intent or the immediate context of use. These could be

collected, however, from complementary methods (e.g., surveys, ethnography).

More innovative naturalistic approaches that leverage smartphone technologies can also be pursued. For example, we [54] used the above methodology to collect data from iPod Touch users. A quasi-experimental design was employed and uncovered differences in usage between socioeconomic status (SES) groups. In particular, we found that lower SES groups used these handheld mobile computers much more and for a wider range of tasks compared to their higher SES peers. This information can be leveraged by designers to accommodate users of different backgrounds (e.g., income level) which, of course, is a central tenet of HF [41]. Other domains could apply a similar approach to other groups of interest (e.g., novice-expert).

Regardless of the topical application, a more naturalistic approach to implementing this technology can better leverage its strengths and uncover real behaviors. These include enhanced access to mediated communications in ecologically valid settings, accuracy in capturing real behaviors, and non-invasive data collection which does not rely on participants or reinforce the fact they are being measured. Other strengths of particular relevance to the study of communications can be inferred as well. For instance, it allows researchers to quantitatively assess the influences of the social, physical, and temporal environment on communications in an integrated way.

6. Conclusion

Clearly, naturalistic studies can be beneficial to the field of communications research to establish empirical patterns and test hypotheses. We argue that researchers in HCI are in a unique position to leverage emerging logging technologies to this end. Many researchers in our field often have the technical background that is necessary for working with the technology as well as the psychological research experience necessary to design, analyze, and apply behavioral data appropriately.

Data gathered from logging methodologies can be useful in understanding communications in ways that standard observational and self-report methodologies cannot. We do not argue that logging should completely replace traditional methodologies. However, we do believe that it is an important method to complement these techniques by providing more accurate, longitudinal, and objective data that cannot be obtained in other ways. The design and implementation of logging studies can be more time consuming and challenging for HCI researchers. However, these enhanced insights into user behaviors can more effectively inform theories, empirical patterns of behaviors, and the next generation of highly usable communications systems.

Acknowledgments

References

[1] ITU, 2010, http://www.itu.int/ITU-D/ict/material/FactsFigures2010.pdf.

[2] R. Ling, *The Mobile Connection*, Elsevier, Dresden, Germany, 2004.

[3] M. Raento, A. Oulasvirta, and N. Eagle, "Smartphones: an emerging tool for social scientists," *Sociological Methods and Research*, vol. 37, no. 3, pp. 426–454, 2009.

[4] D. R. Dubey, R. N. Kent, and S. G. O'Leary, "Reactions of children and teachers to classroom observers: a series of controlled investigations," *Behavior Therapy*, vol. 8, no. 5, pp. 887–897, 1977.

[5] R. E. Sykes, "Toward a theory of observer effect in systematic field observation," *Human Organization*, vol. 37, pp. 148–156, 1978.

[6] K. E. Weick, "Systematic observational methods," in *Handbook of Social Psychology*, G. Lindzey and E. Aronson, Eds., 1968.

[7] F. C. Harris and B. B. Lahey, "Subject reactivity in direct observational assessment: a review and critical analysis," *Clinical Psychology Review*, vol. 2, no. 4, pp. 523–538, 1982.

[8] J. G. Kelly, "Naturalistic observations and theory confirmation: an example," *Human Development*, vol. 10, no. 3, pp. 212–222, 1967.

[9] J. Kjeldskov and C. Graham, "A review of mobile HCI research methods," in *Proceedings of the International Conference on Human-Computer Interaction*, Udine, Italy, 2003.

[10] B. Beurer-Zuellig and M. Meckel, "Smartphones enabling mobile collaboration," in *Proceedings of the 41st Annual Hawaii International Conference on System Sciences (HICSS '08)*, January 2008.

[11] R. E. Grinter and M. A. Eldridge, "Y do tngrs luv 2 txt msg?" in *Proceedings of the Computer Supported Cooperative Work (CSCW '01)*, pp. 219–239, 2001.

[12] N. Eagle, A. Pentland, and D. Lazer, "Inferring friendship network structure by using mobile phone data," *Proceedings of the National Academy of Sciences of the United States of America*, vol. 106, no. 36, pp. 15274–15278, 2009.

[13] P. Hagen, T. Robertson, M. Kan, and K. Sadler, "Emerging research methods for understanding mobile technology use," in *Proceedings of the Australasian Computer-Human Interaction Conference (OZCHI '05)*, Canberra, Australia, 2005.

[14] J. Kjeldskov and J. Stage, "New techniques for usability evaluation of mobile systems," *International Journal of Human Computer Studies*, vol. 60, no. 5-6, pp. 599–620, 2004.

[15] C. D. Wickens and J. G. Hollands, *Engineering Psychology and Human Performance*, Prentice Hall, New Jersey, NJ, USA, 3rd edition, 2000.

[16] J. Rubin, *Handbook of Usability Testing*, John Wiley & Sons, New York, NY, USA, 1994.

[17] J. E. Driskell and E. Salas, "Collective behavior and team performance," *Human Factors*, vol. 34, no. 3, pp. 277–288, 1992.

[18] J. Grudin, "Why groupware applications fail: problems in design and evaluation," *Office: Technology and People*, vol. 4, pp. 245–264, 1988.

[19] R. K. Merton, *Social Theory and Social Structure*, The Free Press, Glencoe, Ill, USA, 1949.

[20] N. Eagle and A. Pentland, "Reality mining: sensing complex social systems," *Personal and Ubiquitous Computing*, vol. 10, no. 4, pp. 255–268, 2006.

[21] S. Carter and J. Mankoff, When participants do the capturing: The role of media in diary studies. *Human-computer interaction institute*, 2005, http://repository.cmu.edu/hcii/123/.

[22] M. Kamvar, M. Kellar, R. Pater, and Y. Xu, "Computers and iPhones and mobile phones, oh my! A logs-based comparison of search users on different devices," in *Proceedings of the WWW: International World Wide Web Conference*, Madrid, Spain, April 2009.

[23] A. Kivi, "Measuring mobile user behavior and service usage: methods, measurement points, and future outlook," in *Proceedings of the 6th Global Mobility Roundtable*, Los Angeles, Calif, USA, 2007.

[24] K. Purcell and K. Brady, "Adaptation to the invasion of privacy: monitoring behavior with a miniature radio transmitter," *Merrill-Palmer Quarterly*, vol. 12, pp. 242–254, 1965.

[25] L. D. Catledge and J. E. Pitkow, "Characterizing browsing strategies in the world wide web," in *Proceedings of the 3rd International World Wide Web Conference*, Darmstadt, Germany, 1995.

[26] B. J. Jansen, A. Spink, and T. Saracevic, "Real life, real users, and real needs: a study and analysis of user queries on the Web," *Information Processing and Management*, vol. 36, no. 2, pp. 207–227, 2000.

[27] L. Tauscher and S. Greenberg, "Revisitation patterns in World Wide Web navigation," in *Proceedings of the 1997 Conference on Human Factors in Computing Systems (CHI '97)*, pp. 399–406, March 1997.

[28] V. Cothey, "A longitudinal study of World Wide Web users' information-searching behavior," *Journal of the American Society for Information Science and Technology*, vol. 53, no. 2, pp. 67–78, 2002.

[29] M. C. Burton and J. B. Walther, "The value of web log data in use-based design and testing," *Journal of Computer-Mediated Communication*, vol. 6, no. 3, 2001.

[30] B. McKenzie and A. Cockburn, "An empirical analysis of web page revisitation," in *Proceedings of the 34th Annual Hawaii International Conference on System Sciences (HICCS '01)*, January 2001.

[31] P. W. Foltz and M. J. Martin, "Automated communication analysis of teams," in *Team Effectiveness in Complex Organizations*, E. Salas, G. F. Goodwin, and S. Burke, Eds., Taylor Francis, New York, NY, USA, 2008.

[32] P. A. Kiekel, N. J. Cooke, P. W. Foltz, and S. M. Shope, Automating, 2001.

[33] C. A. Bowers, F. Jentsch, E. Salas, and C. C. Braun, "Analyzing communication sequences for team training needs assessment," *Human Factors*, vol. 40, no. 4, pp. 672–679, 1998.

[34] A. G. Miklas, K. K. Gollu, K. K. W. Chan, S. Saroiu, K. P. Gummadi, and E. De Lara, "Exploiting social interactions in mobile systems," *Lecture Notes in Computer Science*, vol. 4717, pp. 409–428, 2007.

[35] E. Yoneki, "Visualizing communities and centralities from encounter traces," in *Proceedings of the 3rd ACM Workshop on Challenged Networks (CHANTS '08)*, pp. 129–131, September 2008.

[36] K. Farrahi and D. Gatica-Perez, "What did you do today? Discovering daily routines from large-scale mobile data," in *Proceedings of the 16th ACM International Conference on Multimedia (MM '08)*, pp. 849–852, October 2008.

[37] N. Eagle and A. S. Pentland, "Eigenbehaviors: identifying structure in routine," *Behavioral Ecology and Sociobiology*, vol. 63, no. 7, pp. 1057–1066, 2009.

[38] L. Hossain, K. Chung, and S. Murshed, "Exploring temporal communication through social networks," in *Proceedings of the 11th IFIP TC 13 International Conference on Human-Computer Interaction (Interact '07)*, Springer, 2007.

[39] S. Hulkko, T. Mattelmäki, K. Virtanen, and T. Keinonen, "Mobile probes," in *Proceedings of the 3rd Nordic Conference on Human-Computer Interaction (NordiCHI '04)*, pp. 43–51, Tampere, Finland, October 2004.

[40] A. Oulasvirta, R. Petit, M. Raento, and S. Tiitta, "Interpreting and acting on mobile awareness cues," *Human-Computer Interaction*, vol. 22, no. 1-2, pp. 97–135, 2007.

[41] B. Shneiderman, "Universal usability," *Communications of the ACM*, vol. 43, no. 5, pp. 85–91, 2000.

[42] J. Häkkilä and C. Chatfield, "'It's Like if you opened someone else's letter' user perceived privacy and social practices with SMS communication," in *Proceedings of the 7th International Conference on Human Computer Interaction with Mobile Devices and Services (MobileHCI '05)*, pp. 219–222, September 2005.

[43] M. Raento, A. Oulasvirta, R. Petit, and H. Toivonen, "ContextPhone: a prototyping platform for context-aware mobile applications," *IEEE Pervasive Computing*, vol. 4, no. 2, pp. 51–59, 2005.

[44] S. M. Johnson and O. D. Bolstad, "Reactivity to home observation: a comparison of audio recorded behavior with observers present or absent," *Applied Behavioral Analysis*, vol. 8, pp. 181–185, 1975.

[45] S. N. Haynes and W. F. Horn, "Reactivity in behavioral observation: a review," *Behavioral Assessment*, vol. 4, no. 4, pp. 369–385, 1982.

[46] S. M. Johnson and O. D. Bolstad, "Methodological issues in naturalistic observation: some problems and solutions for field research," in *Behavior Change: Methodology, Concepts, and Practice*, L. A. Hamerlynck, L. C. Handy, and E. J. Mash, Eds., Research, Champaign, Ill, USA, 1973.

[47] P. A. Adler and P. Adler, "Observational techniques," in *Collecting and Interpreting Qualitative Materials*, N. K. Denzin and Y. S. Lincoln, Eds., Sage, Thousand Oaks, Calif, USA, 1988.

[48] C. C. Tossell, P. T. Kortum, C. W. Shepard, A. Rahmati, and L. Zhong, "Assessing the performance of common tasks on handheld mobile computers," in *Proceedings of the 54th Human Factors and Ergonomics Society Annual Meeting (HFES '10)*, pp. 542–546, October 2010.

[49] T. Matthews, J. Pierce, and J. Tang, "No smartphone is an island: the impact of places, situations, and other devices on smartphone use," IBM RJ10452, 2009.

[50] C. C. Tossell, P. Kortum, C. Shepard, L. H. Barg-Walkow, A. Rahmati, and L. Zhong, "A longitudinal study of emoticon use in text messaging from smartphones," *Computers in Human Behavior*, vol. 28, pp. 659–663, 2012.

[51] C. C. Tossell, P. Kortum, A. Rahmati, C. Shepard, and L. Zhong, "Characterizing web use on smartphones," in *Proceedings of the Human Factors in Computing Systems (CHI '12)*, ACM, 2012.

[52] B. Jönsson, A. Svensk, D. Cuartielles, L. Malmborg, and P. Schlaucher, "Mobility and learning environments—engaging people in design of their everyday environments," Tech. Rep., 2002.

[53] P. M. Podsakoff, S. B. MacKenzie, J. Y. Lee, and N. P. Podsakoff, "Common method biases in behavioral research: a critical review of the literature and recommended remedies," *Journal of Applied Psychology*, vol. 88, no. 5, pp. 879–903, 2003.

[54] C. C. Tossell, J. Jardina, P. Kortum et al., "Effects of socioeconomic diversity on iPod touch device use in real-world environments," in *Proceedings of the 56th Annual Meeting of the Human Factors & Ergonomics Society*, Las Vegas, Nev, USA, 2011.