Research In Brief: Using Mobile Phones to Collect Daily Experience Data From College Undergraduates

Russell D. Ravert  Shaun I. Calix  Michael J. Sullivan

This research brief describes our recent efforts collecting daily experience data from college undergraduates at a large midwestern U.S. university through mobile phone text messaging. By daily experience data, we mean data that are collected at multiple points from individuals within their natural context, over a period of time. This approach to data collection provides a way to study phenomena under the conditions in which they naturally occur and to examine how those phenomena progress over time or across contexts (Bolger, Davis, & Rafaeli, 2003).

Gathering data from individuals at multiple points over the course of time as a way to better understand their experiences has been used as early as the 1920s. Since then, the methods of collecting that data have evolved alongside technological advances, with early paper and alarm watches eventually giving way to beepers and personal digital assistant (PDA) devices (Scollon, Kim-Prieto, & Diener, 2003). A variety of approaches are possible, including (a) time-based designs, in which participants are asked to respond at fixed intervals (e.g., at 10:00 a.m. and 4:00 p.m. each day), (b) event-based designs, in which participants are asked to respond when a certain event occurs (e.g., before each meal), and (c) interval-based designs, in which participants are asked to respond whenever prompted (e.g., by an electronic beeper). In the current study, we examined the feasibility of an interval-based approach that might be considered a natural extension of these methods: collecting data from college students via text messaging.

The Experience Sampling Method

The experience sampling method (ESM) is a term associated with interval-based designs in which participants provide daily experience data when they are signaled at various (usually random) times during the day and across an extended period of time (Hektner, Schmidt, & Csikszentmihaly, 2007; Scollon et al., 2003). Other compatible terms include ecological momentary assessment (Stone, Shiffman, & DeVries, 1999) and time-based diary research (Bolger et al., 2003). Collecting data through this method has several strengths compared with traditional survey or laboratory-based methodologies. First, experience-sampling allows a useful way to explore the link between context and behavior or feelings, because data can be collected while the participant is within a particular context. Second, time-based methods allow the ability to assess changes that occur within individuals over time or across situations. Third, the accuracy of data need not rely on participants’ retrospective memory, as is often required in traditional survey methods.

A large number of ESM studies have focused on adolescence. Topics of ESM...
research conducted with that population have included studies of time use (Larson, 1989), the context of mood (Larson, Moneta, Richards, & Wilson, 2002), student engagement during instructional activities (Shernoff, Csikszentmihalyi, Schneider, & Shernoff, 2003), and the relationship of cortisol levels to emotions (Adam, 2006). ESM methods have also been utilized to better understand the experiences of college students. For example, in order to study the experiences of Black students on predominately White campuses, Cole and Yip (2008) provided Black college freshmen with electronic pagers and paper data diaries. Over a 10-day period, participants logged data regarding their location and mood whenever they were beeped. The multiple data points allowed the researchers to explore the relationship between participants’ emotional states in school versus nonschool settings. Other research topics studied in college settings through daily experience data methods have included motivators of alcohol use (Hussong, 2003), and events evoking social anxiety (M. R. Lee, Okazaki, & Yoo, 2006). In their study of risk perceptions among college students, Hogarth, Portell, and Cuxart (2007) used an event-based ESM design, but utilized students’ own mobile phones rather than providing pagers. In that study, participants were supplied with questionnaires that they were asked to complete whenever they received a text message from the researcher.

Researchers are not limited to having participants complete responses on paper, however. In a number of EMS studies, participants have been provided with PDA devices so that they could enter responses directly into those devices whenever prompted (for a review of methods including PDA use, see Hektner et al., 2007). Each approach appears to come with advantages and disadvantages. In reviewing electronic and paper-based methods, Broderick (2008) pointed out that one limitation to paper-based methods is that they assume that respondents will complete the form at the time of being signaled. That author’s attempts to verify compliance with paper-based methods found evidence that participants frequently completed the requested data before or after the actual time of prompting. In comparing paper-based and electronic methods, Broderick concluded that providing PDA or other devices may hold advantages in terms of compliance, less missing data, and fewer data-entry errors; but doing so may hold disadvantages including technical problems, cost, and the need for training participants.

Text Messaging as a Research and Intervention Tool

As mobile phone use and text messaging has continued to play a more central role in people’s daily lives, some researchers and clinicians have sought to use the medium as a tool for interventions and data collection. Such efforts have included sending tailored health messages to college students who are trying to quit smoking (Obermayer, Riley, Asif, & Jean-Mary, 2004), adolescents with diabetes (Franklin, Waller, Pagliari, & Greene, 2006), adults in a weight-loss program (Patrick et al., 2009), and patients with eating disorders (Bauer, Percevic, Okon, Meermann, & Kordy, 2003).

In one of the few known published studies that utilized text messages to systematically collect diary data, asthma patients (median age = 38.5 years) were asked to reply to at least three of four messages sent daily for 2 months (Anhøj & Møldrup, 2004). Message content included a variety of self-management topics and data. The researchers found less attrition than they had experienced in Web-based diary efforts and concluded that the method was feasible and resulted in acceptable response rates.
Text messaging appears to have several strengths as a tool to collect data from college students regarding their daily experiences. Like other ESM methods, data collection via mobile phone would allow data collection under natural environmental conditions and across multiple time points. But because text messaging is especially prevalent among youth (Faulkner & Culwin, 2005), using that medium for data collection might be a more viable alternative to implement than providing PDA or other electronic devices. In a survey of undergraduates \( N = 250 \) conducted at the university where the current study took place, over 98% of students reported owning a mobile phone, and 85.6% of those students reported using text messaging (Ravert & Kile, 2007). Having participants use their own mobile phones would reduce costs and security risks associated with providing participants with PDAs or other electronic devices (Hektner et al., 2007). An additional advantage might be that, because mobile phones are already a part of college culture, students might be more likely to respond to messages in situations where they might be inclined to forgo paper or PDA methods (e.g., at a bar or beach).

The Current Study

We undertook the current pilot study to explore the feasibility of collecting quantitative and qualitative data from undergraduates by sending them text messages via their mobile phone. Our method involved asking participants to agree to receive and reply (within 15 minutes) to a series of daily messages sent over a 1-week period. Our objective was to assess the general feasibility of text messaging as a data collection tool and to identify problems or issues that might jeopardize reliability or validity. We were especially interested in evaluating (a) what response rate the approach would achieve, (b) how well text messaging could facilitate collection of both quantitative and qualitative data, (c) what length and legibility of replies would result, (d) how promptly participants would reply to the daily messages, and (e) how users would perceive the method. In this brief we report findings and discuss the challenges and potential of the approach for researchers and student affair practitioners who might consider using similar methodologies to collect data on college student behaviors, attitudes, and experiences.

METHOD

Participants

Sixteen undergraduates, ages 19-23 \( M = 20.4 \) yrs), were recruited from a large (approximately 200 students) undergraduate human development and family studies course. Procedures used in the study were approved by the campus institutional review board. Following a course announcement, we randomly selected a set of potential participants from the course roster and sent those students a recruitment e-mail. Students responding to the e-mail were invited to schedule a time to visit the researcher’s lab for a brief intake session. A randomly selected study ID was assigned to each participant and used to replace identifying information (including phone number) in the final dataset in order to protect confidentiality.

The final sample included 14 females and 2 males and was evenly distributed among sophomores, juniors, and seniors. A large majority (87.5%) of students indicated White as their race. The demographic makeup of the sample was representative of the class, but was not representative of the university. Still, we considered the sample acceptable due to the exploratory nature of the study and given that our goal was to assess feasibility of a method (rather than to generate generalizable inferences).
Procedure

At the intake sessions, participants reviewed and agreed to the study consent form, were informed about the study, and completed a Web-based intake survey. We explained that they would receive the same text message twice a day, worded, “What are you doing now? How risky is what you are doing, 0-9? What could happen?” We requested that replies always include (a) a brief, legible sentence describing the activity they were engaged in when they received the text message; (b) a number, using a scale from 0 (none) to 9 (extremely risky), to describe how risky they considered their current activity; and (c) a brief explanation of what outcome could result from the current situation (if the rating had been anything other than 0).

Participants completed surveys that collected demographic data including the students’ age, year in school, gender, and race. In addition, they were asked to provide a list of times they would not be available to receive the text messages (i.e., during class). Although we were unable to tailor message times to fit individual schedules, we explained that messages would not be sent at times when a majority of participants were unavailable. Participants were told what date the messages would begin, but were not given the specific time that messages would be sent.

Participants received incentives of $10 for completing the intake session, $1 for every text message they replied to (14 possible), and $15 for successfully completing the study and Web-based exit survey. They also received 10 cents per text message received or sent to reimburse any charges they might have accrued.

Data Collection

The research team consisted of the primary investigator, a doctoral-level research assistant, and a university information technology administrator. Text messages were sent from and received by a university e-mail account created for use in the study. Two methods of text message delivery were piloted. The first set of messages was sent manually as a single message, using Microsoft Outlook, by including participant addresses in the blind carbon copy (bcc) field. The second set of messages was sent using a Visual Basic Script, developed by the third author, that automatically sent the set of messages at predetermined times. The ability to send messages at any predetermined time during a 24-hour cycle was one of the main benefits of using the automated system.

Participants were simultaneously sent a text message twice per day, for 7 days. The exact times of the text messages varied each day and were unknown to the participants. One message was sent between 9:00 a.m. and 4:00 p.m. daily, and the second was sent between 5:00 p.m. and 10:30 p.m. Message times were varied as widely as possible across the 7 days, purposely avoiding times that a large number of participants had indicated as being unavailable. For example, no messages were sent on weekdays between 9:00 a.m. and 11:00 a.m. because a large number of students had listed that time as unavailable (due to classes). All outgoing messages were worded, “What are you doing now? How risky is what you are doing, 0-9? What could happen?” This choice of wording allowed us to assess the feasibility of collecting quantitative as well as qualitative data.

Results

Participation and Message Reception

In intake sessions we found that, although students all knew how to receive a text message on their mobile phone, none knew what e-mail address we would need to use in order to send them text messages from a computer. The standard protocol is that text
messages are addressed by the user’s mobile phone number followed by @ and an extension specific to the cellular service provider for that individual (e.g. 5555555555@cingularme.com). Because participants did not know what extension should be used, we generated an initial e-mail address based on the participant’s mobile phone number and the extension that we thought to be required by their respective service providers. In one case, a participant’s first message was returned and we were successful after trying an alternate extension for the same service provider.

Two students who signed up to participate had to be dropped from the study due to technical problems that kept them from receiving our text messages. In one case, the participant sent the researchers a message on the second day of data collection explaining that she/he received notification that a message had been delivered, but did not receive any message content. In the second case, a participant sent the researchers an e-mail on the fourth day of the study saying that she/ he hadn’t received any text messages yet. We attempted to discuss the problem with the service provider, but their policy allowed only talking directly with the client, and we were ultimately unable to uncover the cause of the problem. The two students unable to complete the study were provided with incentives for their participation but are not included in analysis.

Response Rate

Among the 14 participants (excluding the 2 who were dropped due to technical problems), 4 students replied to all 14 messages for a 100% response rate. The lowest response rate was 64.3% for a participant who responded to 9 of 14 possible messages. Across participants, a total of 174 responses were received (out of 196 possible), representing an overall response rate of 88.8%.

Quality and Length of Responses

The 174 responses we received were examined to assess how frequently they included answers to all 3 questions included in the outgoing message (i.e., what the participant was currently doing, a numerical rating of how risky that activity was, why the activity was risky). All 174 responses were found to describe what the participant was doing (i.e., “watching tv,” “i’m at the gym spotting someone,” and “I’m studying in my room”). In 97.7% of those cases, the reply also included a numerical rating of risk associated with that activity. Those ratings ranged from 0 (i.e., for “sleeping” and “sitting on my couch”) to 8 (i.e., for “texting while driving” and “using a knife”), with a mean rating of 1.73 (SD = 2.01). The final component, why the activity was risky, was addressed in 91.9% of responses (i.e., “I could get in an accident,” “I could screw up and get a bad grade,” and “nothing really could happen”).

Participants typed an average of 42.6 characters (not including spaces), or 10.6 words, per reply. The shortest reply was 16 characters, worded, “Driving 7 accident.” The longest reply included 109 characters (or, 139 characters including spaces) and was worded, “i just drove home from campus level 8 the streets are very dangerous and an accident is likely or falling on the ice getting to the car.” As a point of reference, the maximum length for text messages is typically 160 characters, including spaces. All messages were legible, and no text message short hand or other nonstandard jargon was included in messages.

Promptness of Replies

One purpose of the study was to assess how promptly participants would respond to the daily messages. Across all 174 replies, a large majority (81.8%) arrived in the researcher’s
inbox within 15 minutes of the outgoing message. The mean time of reply (from the time the message was sent until a reply was received) was 14.5 minutes, with a median of 4 minutes. That mean was skewed by a small number of very slow responses. Four messages (2.3% of replies) were received more than 1 hour following the outgoing message. The longest delay in replying was 10 hours, for a message saying, “I was sleeping when u text last night 0 risk nothing could have happened.” In some cases the cause of a late response was evident within the message, as in the case of one late message that explained that the participant had “just left a test.” At other times it was not clear why a participant had failed to reply within the requested 15-minute limit. For example, one message worded “I am at a party. Rate 5 i could get drunk & do something stupid” was received 90 minutes after the outgoing message was sent.

**User Perceptions**

In an anonymous Web-based survey conducted upon completion of the week of text messaging, participants were asked what, if anything, had kept them from responding to all of our messages. Their explanations included that they (a) were sleeping ($n = 3$), (b) had the phone on silent ($n = 2$), (c) were working ($n = 2$), (d) were not near the phone ($n = 1$), and (e) received the message during an exam ($n = 1$).

Asked how accurate their responses were, 57.1% of participants indicated “extremely accurate,” and 42.9% selected “quite accurate.” No participants selected “slightly accurate” or “not at all accurate.” Asked why their message might not have been 100% true, one participant explained that some time had passed before she/he replied, possibly limiting the accuracy of the response. Another participant described “feeling rushed” while responding and a third replied that her/his responses depended on what mood she/he was in at the time. One participant reported giving “quite accurate” replies in general, but added that in some cases she/he might have embellished because she/he “wasn’t doing anything” when the message came. The participant wrote, “I had to make it somewhat interesting, so it may not have been all true.”

When asked their opinions about the questions that had been included in the daily text message they received, participants expressed general satisfaction. All participants described the first question (“what are you doing now?”) favorably. However, there were different opinions regarding the quantitative scale used in the second question (“How risky is what you are doing, 0-9?”). Whereas some participants described the scale as “easy to understand” and “straight forward,” one participant described it as “a little tricky.” A second participant wrote that the scale was “hard to judge,” and that “the degree of ‘0’ and ‘9’ should be more clear.” Thus, it may be important to include the rating scale within the message itself, in addition to clarifying it at the beginning of the study as we did.

**Limitations to the Method**

Although we were pleased with the success of the text message data collection method overall, several limitations to the methodology were noted. Some of these limitations involve our decision to send and manage text messages within an e-mail environment. As noted, we were unable to identify an e-mail address that would successfully send a text message to the mobile phones of two participants. Also, we found that our e-mail program (Microsoft Outlook) did not store all of the data we wished regarding when and where messages were sent. For example, when messages were sent using the blind carbon copy (bcc) function, no record was automatically kept regarding which addresses the message was sent to. Further,
when we began retrieving participant’s text replies, we found (unexpectedly) that some of those messages had been automatically forwarded to the junk mail box.

These challenges related to sending and receiving text messages through a computer-based e-mail program might be addressed by sending messages from a custom application or from another mobile phone. But other limitations involve the unique medium of text messaging. The first regards a lack of available information regarding the exact time that participants actually received and sent messages. Text messaging is unique compared with telephone calls, beepers, or PDAs, in that message transmission cannot be assumed to occur instantly. In cases where no reply is received, the researcher had no certainty that the participant received the message. Additionally, text messages are subject to being placed in cues or being rerouted before reaching the receiver. Although some service providers automatically include the time of reception in reply messages, that function is not standard. Therefore, the exact time that participants received the message was not always available. Likewise, the exact time that they sent a reply was not always available; only when we received that reply was known. So, identifying the amount of time that passed between the time a participant received and replied to a message could not be accurately assessed. Rather, we were limited to knowing when the original message was sent and when the reply was received. This limitation could be a substantial problem in studies where validity of data is strictly time dependent. Also, our ability to make participant incentives contingent on timeliness of replies was limited by a lack of accurate information regarding how much time might have passed before participants replied. Future feasibility studies of text messaging might use an alternate means to assess the speed and reliability of messaging from a technical perspective, for example by having participants call to notify the researcher of the exact time that they receive and reply to the messages.

Although the data we collected was confidential, and responses were separated from identifying information, anonymity was not an option in the study because we were sending messages to participant’s personal cell phones. The consent form specified that although responses would be kept confidential, under certain, extremely rare circumstances (for example, if presented with a court order to divulge the records), we might be required to divulge content of their responses to a third party. Although we had no indication that confidentiality concerns influenced results in any way, those types of concerns could potentially limit participant’s willingness to fully disclose their actions in responses.

Also, for some research objectives, the limits (and norms) associated with the length of text messages might be problematic. Although replies we received were consistently legible and typically addressed all components of the question we had asked, they often lacked contextual information and richness that would be possible in other methods such as surveys or interviews. For example, one participant simply responded, “working 3 could get hurt,” without including details of the type of work or what injury she/he expected might occur. For some research questions, brief responses such as this one might be sufficient. But in other cases, the researcher might desire a degree of detail and elaboration that exceeds the capacity allowed in text messages (160 characters). Additionally, students who are used to sending brief, single sentence text messages might be inclined to do so when responding to research questions as well. Researchers might consider using additional methods (i.e., e-mail, phone interview) to follow-up on data that is collected via text messaging but needs further explanation.
RECOMMENDATIONS

Text messaging was found to hold substantial potential as a way to collect small amounts of qualitative and quantitative data across multiple time points from college students. The response rate was exceptionally high. In a large majority of cases, participants responded as requested, providing legible text and a numerical rating. Our experience suggested several recommendations for others planning to use this method. The first set of suggestions includes issues to be addressed prior to data collection.

In presenting the study protocol, researchers should be clear about expectations in replies—for example, whether or not text shorthand or jargon is acceptable. Researchers might want to stress that participants provide an accurate and honest answer to each message, regardless of how well they think it matches the researchers’ intentions (for example, even if they feel what they are writing is boring).

From a safety and human subjects perspective, we were concerned to see that almost one half of participants reported replying to our message while driving, even after having been specifically instructed not to do so and signing a consent form agreeing not to do so. The dangerous practice of using text messaging and other distractive technologies while driving is thought to be especially problematic among young drivers (J. D. Lee, 2007). For this reason, researchers might stress the importance of avoiding texting while driving or other hazardous activities.

Participation seemed very manageable for this sample of college students. Few inconveniences or suggestions were noted by participants. The most common barrier to replying that participants mentioned was receiving messages at a time when they were unavailable due to sleep, work, or class, or when the phone was on silent. However, technical problems kept two potential participants from participating, which could be a substantial limitation in terms of representativeness. Testing the text address during the intake session would be helpful by identifying whether problems with delivery exist. In that case, the participant might call the service provider with the researcher present, so that they can discuss the problem with a service provider representative to identify the problem and find a solution. Sending an e-mail to all participants after the first day of messages, confirming the number of messages they received and how many replies they sent, might also be useful.

Researchers might prepare for some degree of inevitable uncertainty regarding exactly when participants received the outgoing message or sent a reply. Also, researchers should recognize that they have no assurance that a text message sent to multiple recipients simultaneously will be received at the same time by all recipients. Studies that depend on precise timing of responses might not be best suited for text messaging due to these limitations.

Two additional recommendations based on our experience involve organization and management of data when sending messages from an e-mail program. First, researchers using the method might consider manually logging the time and recipient of all outgoing messages, because that information may not be automatically stored. Also, our experience with finding replies in the junk mail folder suggests the importance of checking that message replies and checking settings to insure that any message inadvertently forwarded to that location is not automatically deleted.

CONCLUSION

Text messaging as a data collection method appears to hold substantial potential for
Research in Brief

collecting numerical and text responses, but within a set of limitations. When standardized responses are required or certainty regarding response timing is required, PDA devices or similar methods may have an advantage over text messaging (for example, we have recently been piloting collecting ESM data from college students using automated push-button telephone surveys). But if that is not the case, and when researchers wish to collect small amounts of open-ended responses across multiple time points, text messaging may be a good solution. Advantages to the text message data collection method appear to be low cost, minimal time required for training and implementation, favorable response rates, and that the existing popularity of text messaging minimizes the need to provide participants with equipment. Limitations appear to include the potential for technical problems, uncertainty regarding message reception, and restrictions on the amount of text that can be sent and received per message.

In a review of experience sampling studies, Scollon, and colleagues (2003) concluded that the method “is most useful when applied in conjunction with other methods” (p. 12). Based on the strengths and limitations that we identified, we feel that using text messaging as a data collection tool might be most effective in studies that use the method along with other strategies (i.e., surveys, observations, or interviews). One approach would be to use follow-up interviews, daily e-mails, or other tactics to illicit elaboration on data that were collected in the text messages.

In conclusion, collecting data from students within the context of their natural environment via experience sampling methodology appears to be a valuable methodology for researchers and student affairs professionals who seek to collect data from students regarding their day-to-day experiences. Text messaging appears to hold promise as a tool for doing so. We hope that researchers and practitioners will be able to capitalize on strengths of the method as a means to improve the college experience and to better understand college student behavior and development.

Correspondence concerning this article should be addressed to Russell D. Ravert, Department of Human Development & Family Studies, 314 Gentry, University of Missouri, Columbia, MO 65211; ravertr@missouri.edu
REFERENCES


