WAT R U DOIN? Studying the Thumb Generation Using Text Messaging[1]

Eszter Hargittai and Chris Karr
Northwestern University

A unique opportunity

The method of data collection that relies on calling respondents has suffered from declining response rates for years. Concurrently, a new medium of communication – text messaging using one's cell phone – has gained popularity recently, especially among younger populations. The goal of this study was two fold: methodological and substantive. We wanted to see whether we could collect time diary data using text messaging as our data collection tool from college students about their everyday activities. In particular, we wanted to focus on communication processes and exposure to digital media so in addition to collecting traditional time-diary data, we wanted to make sure to obtain information about participants' communication practices during the day.

We were piggybacking on a larger study already under way with questions that we believed merited an exploration even if, in the end, our time-diary data collection would turn out to be exploratory at best. Overall, we had both methodological and substantive reasons to pursue the work and given what was already being invested in the larger project, we decided that the marginal costs were worth our time and effort. In hindsight, we are very happy to have pursued this opportunity and although it took more time than we anticipated, we have achieved some very interesting and unique results that were well worth the investment.

Hargittai was in the midst of working on a two-year project studying adolescents' Internet uses, skills and participation using surveys and in-person observations. The study

[1] The authors are grateful to Viktor Domazet and Alex Knell for their helpful input. They also acknowledge the assistance of Elizabeth Anderson, Waleeta Canon, Dan Li and Gina Walejko. Many thanks go to the National Science Foundation (IIS0712874), the John D. and Catherine T. MacArthur Foundation, and the Northwestern University School of Communication for their support. Hargittai is grateful to the Center for Advanced Study in the Behavioral Sciences and The Lenore Annenberg and Wallis Annenberg Fellowship in Communication for time to work on this project.

also had a longitudinal component whereby some participants would be randomly assigned into a training program and at a later point in time would be observed again – along with those who had not received training – to test whether the intervention had made a difference in students' online know-how. Given these parameters, it was already a given that some students would be approached for participation more than once. More importantly, we were already collecting data on respondents that could then be merged with additional information we were hoping to collect about them using text messaging.

Every methodology has its limitations and we can only learn so much about any topic using just one method. One challenge of surveys – the main method of data collection in the larger study – is that it is hard to gather nuanced and reliable information about the details of people's everyday time uses. This concern prompted the idea of trying to gather some additional time diary data from respondents. However, given students' busy lives and the difficulty in convincing people to participate in recurring studies, the challenge remained: How to collect diary data from mostly 18 year olds who are physically hard to pin down amidst their busy college lives?

This age group is sometimes referred to as the Thumb Generation, because young adults spend so much time on their cell phones both for calling people in their networks, but also texting them using the dial pad of their phones. Data collected by Hargittai one year earlier about a similar group suggested that most students in the population of the study owned cell phones and many used text messaging, so the idea of this being a popular activity was not simply based on unsubstantiated assumptions, rather, systematic data. In fact, a look at the survey responses of the current study's sample made it clear that over 98 percent owned a cell phone of whom over 90 percent used the device for text messaging. This prompted the idea that we should collect diary data through the relatively unobtrusive medium of text messaging, a method that does not require physical copresence between researchers and respondents, and draws on an activity in which students are already engaged during their everyday lives.

In what follows, we will give a bit more detail about the methodological and substantive motivations for the study.[2] We will then say a few words about some distinct

[2] Because this chapter is being written while the study is still under way, this piece reports on the initial phases. The bulk of methodological issues have been addressed by the time of this writing. The ongoing parts concern additional waves of data collection replicating the methods described herein.

features of this collaboration. Next, we offer a full timeline of the research project for a realistic overview of what type of time commitment a study of this sort entails. One's first reaction may be that collecting diary data through text messaging should be fairly simple – that is what we had thought! –, but not surprisingly, just like with any other research project, once one hits the ground, complexities emerge from every direction. Having described the motivation and context of the study, we offer detailed descriptions of the following important components: establishing and setting up the technical and logistical system for sending and receiving text messages, developing and revising a coding scheme, building and refining the coding interface, and finally, collecting the data. We conclude with a discussion of main lessons learned and what challenges may be encountered when trying to scale up from our experiences.

Long-distance interdisciplinary collaboration

While somewhat tangential to the study, an important point we want to get across is that long-distance collaboration is very much feasible in this day and age, so a lack of physical copresence should not deter researchers from pursuing a joint project. The lead on the overall study was a junior faculty member (Hargittai) on sabbatical from Northwestern University at the Center for Advanced Study in the Behavioral Sciences in Stanford, California. The collaborator on this study (Karr) was in his first year of graduate school in the Media, Technology and Society Ph.D. program at Northwestern University.

This work cut across academic positions (faculty/student), disciplines (communication/sociology/computer science/psychology) and distances (California/Illinois). Hargittai approached Karr to see if he had an interest in the study and, after a positive response, the collaboration began. We established early on that outcomes would be co-authored assuming similar levels of input from both researchers. Given some of the technical details involved with the project, the interdisciplinary nature of the partnership worked to our advantage. It posed some challenges especially when communicating certain ideas given the different terminologies used in various disciplines (see also Sandvig's piece in this volume). However, frequent communication – mainly using email – helped clarify any confusion in a timely manner. The upside of such group effort is not only that different types of tasks can be addressed quickly internally by team members (i.e., it is not necessary to hire a programmer if a tool needs to be developed), but also that

the researchers are very likely to learn about new concepts, terms and tools associated with the work.

Because Hargittai was on sabbatical two thousand miles away from Karr and the study location, almost all of the work on this project happened without any in-person meetings. Given this, the experiences described herein are not simply an example of a lesson in collaboration, but also of long-distance communication and coordination of multiple people and project components. We want to acknowledge the important role that free online services, such as the video-conferencing tool Skype, play in making such undertakings possible.

Finally, while only two names appear on the by-line of this chapter, it is important to note the helpful input from our larger research team members throughout the study. Such expressions of gratitude are usually left for the Acknowledgements section of a paper, but we consider it an important part of our entire research process worth mentioning in a behind-the-scenes piece of this sort. During the time of this study, the research group met weekly to share progress reports and address questions raised by current project specifics. Consequently, the work benefited from all team members' regular feedback. Moreover, much of the coding was done by undergraduate research assistants whose continuous input was very helpful to the project. We address some logistical specifics related to this later in the chapter.

Timeline

Before launching into a detailed description of how we approached the various parts of the project – from figuring out the technical specifications of our messaging system to recruiting respondents, doing the data collection and compiling our coding scheme– we present a timeline of the project. The goal of this is to give the reader a realistic sense of the many behind-the-scenes activities that are an integral part of such a study, but ones that rarely ever see the light of day in publications.

	Nov	Dec	Jan	Feb	Mar	Apr	May	Jun	Jul
Project idea, first email exchanges									
Seeking funding									
IRB (supplement to main study)									
Securing funding									
Taking notes on our methods									
Test of first system (just the co-authors)									
Building our messaging system									
Building our coding interface									
Fine-tuning our coding interface									
First pre-test on research group									
Constructing and fine-tuning our coding scheme									
Training coders									
Coding of pretest data (to test coding interface and scheme)									
Second pre-test on research group (expanded group)									
Recruitment of respondents									
First wave of data collection (15 participants)									
Second wave of data collection (20 participants)									
Third wave of data collection (21 participants)									
Fourth wave of data collection (4 participants)									
Compiling full data set									
Preliminary coding of study data (to test coding scheme)									
Coding of data									
Write-up of methods									

Initial plan of action

As noted earlier, we approached this project with the belief that it would be relatively straightforward. After all, text messaging is a common activity among college students and the technology seems fairly simple. How hard could all this be? Those are, of course, famous last words at the initial stage of any project where the researcher thinks the study in question will be a quick and easy undertaking.

To tackle the methodological issues raised by relying on text-messaging, we planned to send text message requests to which respondents would send replies. As the section below on setting up our system attests, while not impossibly difficult, the process was nowhere near as simple as one might think.

To address the substantive questions, we were interested in collecting four types of data from respondents for each moment in their day when we prompted them for a response:

1. Location: Where is the respondent located?
2. Activity: What is the respondent doing? (multiple activities are possible)
3. Social surroundings: What is the gender and number – if any – of the people with the respondent, and what is their relationship to the respondent?
4. Communication processes: What – if any – communication processes is the respondent engaged in? In particular, is the respondent using any communication media?

This information is in line with episode data collected in traditional time diary studies (Pentland et al. 1999, p.27.). However, the difference here is that we have a particular focus on communication processes and digital media uses. Moreover, our respondents are constrained by the limitations of this medium. Is it realistic to expect such detailed information from respondents in 140-160 characters, which is the limit put on these messages by providers? Although the purpose of this chapter is not to discuss our substantive findings, the results are encouraging. Many respondents shared considerable amounts of information about their whereabouts, allowing us to supplement our survey data with additional details about the role of digital media in their lives. However, as our notes below will attest, gathering this type of information from participants in such restricted form requires communicating detailed instructions to them, which raises some logistical issues.

Setting up the system for sending and receiving messages

We had to keep several issues in mind while considering various technical solutions to our data collection challenge. Our requests were to be received by, and responses sent from, respondents' mobile phones. Our "pinging" system – as we called it, drawing upon the practice of sending short messages to networked machines to assess their availability – would have to meet two main requirements. First, we had to automate the process of

sending out requests for response to participants. That is, we needed the ability to schedule the requests ahead of time and execute that schedule with minimal human intervention. After all, it is not reasonable for any one person to sit next to a machine and send out requests to numerous respondents every hour for a full day, and it is certainly not a very scalable solution if we were to want to run the study – as was our plan – multiple times in the future. Second, we needed a way to collect responses and store them for later aggregation and coding. Given the various issues that may arise during the study, our system needed to be flexible and extensible so that we could modify it to meet our particular needs. This meant exploring and evaluating competing systems to determine the best fit for our project.

Evaluating existing systems for data collection through text-messaging

It is usually best to avoid reinventing the wheel when it comes to various components of a project. Thus, we started by examining a few research systems targeted to studies similar to our own in the hopes that we may be able to use them for our purposes. In the end, this process did not yield any systems that we could adopt for our study, but it helped us clarify the needs of our project. We include this part of the process here, because it is precisely the kind of detail one never sees in write-ups of studies, but it is the type of activity that takes considerable time and effort and therefore must be part of any realistic research plan.

After performing a literature review on related studies, we found a few preexisting systems. We first looked at Momento (Carter, Mankoff and Heer 2007) to determine if this toolkit for ubiquitous computing experiments fit our needs. We evaluated its architecture, documentation, and design goals. Its focus on using SMS messaging to communicate with respondents mirrored our own, but we found that it was a poor fit for our needs. Its origins in the human-computer interaction field limited its applicability by focusing more on being a tool for simulating interactions on mobile devices than being a robust and extensible data collection instrument. Momento is an impressive tool for testing and designing mobile software and devices, but we determined that in order to adapt it to our purposes, substantial additional development would be required. Furthermore, Momento's design required much more human intervention to conduct studies than we could provide.

We also looked at the Experience Sampling Program (Barrett Feldman and Barrett 2001), a toolkit for creating sophisticated time-diary studies like the one we were designing.

Its study setup and analysis features would have been a good match for our project, but it required that the participants be equipped with customized preprogrammed handheld computers running the Palm or Windows Mobile operating systems. Since one of the most novel components of our study was that we were relying on devices participants already owned, requiring specific devices or programs was a significant deterrent. We did not want to be hindered by requirements of specific hardware or software specifications beyond what would be available to anyone who has a text-messaging subscription on a regular cell phone plan. After all, providing study respondents with such devices is cost-prohibitive and introducing a new device recreates many of the problems encountered when doing traditional paper-based diary studies (see, e.g.,Christensen and Feldman Barret 2003; Pentland et al. 1999). Respondents would be required to integrate the new device into their daily routine (decreasing the likelihood that the equipment is continuously present), and we would have to retrieve the apparatus at the conclusion of the study.

We also evaluated several commercial services for sending and receiving messages, but we either found the services too limited, unreliable, or expensive (or a combination of all three) for our needs. We investigated a number of other providers that supply services that avoid problems such as spam filtering, but these services were both too expensive and a poor fit for our needs. They were primarily designed for regular marketing campaigns that either broadcasted one-way announcements or solicited simple responses ("text '1' to vote for the first contestant"). Their pricing models also assumed a longer continuing business relationship that was incompatible with our study's timeline. Furthermore, throughout this process of finding the right solution, we stayed conscious of the need to avoid becoming locked in to any single service provider. This was important in case it disappeared, started charging too much or changed the system in ways that would make it difficult to use for our purposes. Following our strategy, we could switch services without too much setback (both in terms of labor required to update our system and time lost to revising our course of action).

After several weeks of research and investigation, we decided that the best approach would be to create our own system. We had clear goals for it and this greatly assisted in defining both the scope of the project and the necessary features.

Creating our own system

From the beginning, we decided to focus on a simple, yet extensible system that would allow us to develop and deploy it early so that we could take advantage of a rapid iterative testing and development process. We decided to run the test using AOL's instant messenger network since it included free SMS integration allowing users to send short messages to mobile phones. We discovered this feature in our prior day-to-day use of the service. Combined with robust open source libraries that provided access to the AOL network, this was a vital component of our early development and testing. We simulated the study using this system to determine how the software might function in practice. We sent and received messages manually using a compatible instant messaging client.

We asked members of our research group to act as pretesters. Of course, we did not require these team members to participate, but given that most other people in the research group were involved in doing studies on this same sample, and given that we have a collaborative atmosphere in the lab, most research group members willingly participated, giving us helpful feedback. In this simulation, we experimented with the format of the requests and we evaluated the frequency and content of the responses. By format, we mean the phrasing and structure of the requests that fit within the constrained 160 character limit of text messages. Regarding content of the responses, we were curious to see whether we could make sense of the responses and whether they included the type of information we were seeking. As to the frequency of responses, we wanted to know how realistic it was to contact people every hour for feedback about what they were doing. We also solicited input from the respondents to gauge how intrusive and demanding the study was from their perspective. We found that participating in the simulation was not excessively laborious and that we were able to collect the kinds of information that we wanted. With this helpful experience under our belt, we proceeded to build the actual system to be used for the study.

The first focused on the immediate task of data collection. We needed an automatic system that would send requests and collect responses for later analysis. To fill this need, we constructed a Java Web application that maintained a schedule of requests to be sent out at pre-defined times. We created a simple Web-based interface that allowed the manual

scheduling of requests, but we also provided a remote application programming interface[3] to be used by external scripts to batch-schedule complete studies. These scripts were typically less than 100 lines of code, and the bulk of that consisted of listing the respondents and their schedules. To further simplify the implementation, we avoided using a relational database server and used a simpler XML file. We chose XML as the storage format since it is an open text-based standard that may be read and manipulated using a wide variety of programming languages and tools. This provided us the widest latitude for the future creation of tools to parse, translate, and manipulate the collected data. Since our application only required a single standalone software package, we were able to set it up and host it on a departmental server with very little assistance from the local IT staff. Overall, building this component of the system took less than three weeks of part-time effort. This quick development cycle allowed us to conduct fully functional – i.e., automated as opposed to manual compared to our earlier pretest – live tests with lab members and to begin investigating methods for sending and receiving text messages.

As we developed the Web application, we used the AOL network again for the initial testing of the custom-built software. Since we worried that we would be banned for abusing the network if we sent out too many text messages (AOL provides the service for free, and pays the SMS costs on behalf of its users), we began researching suitable replacements for it. A number of commercial firms offered similar text-messaging services, but we found that the services were either too constrained or too expensive for our needs. Many services provided message deliveries, but were unable to receive any responses. Firms that provided both the sending and receiving services charged high setup fees for establishing the necessary mobile presence and short code – the five-digit number used to contact systems through text-messaging – in addition to charging substantial rates for the continued service and maintenance. We found a provider that offered the features we needed for a reasonable per-message fee, but after testing, we realized it was not sufficiently reliable for our study. We discovered that the way the provider sent our messages activated the spam defenses on the mobile phone networks.

---

[3] An application programming interface (API) defines the set of services that a software component provides to other applications and systems. Software developers create and document APIs so that others may use the services with their own software projects. In this case, we defined an API so that others could write their own scripts for scheduling studies using their own preferred programming languages and environments.

In the end, we addressed the various issues by creating an in-house solution that used one of our team member's own mobile phone to send and receive messages.[4] We connected this phone to a recycled lab computer that received commands to send and collect messages and communicated with the phone using open-source programs. We had no problem developing a suitable plug-in that connected to this system. We had some initial concerns about the reliability and cost of using a regular mobile phone for this purpose, but this ended up being more reliable than any of the alternatives we previously considered. Furthermore, it was less costly than the other options, even though we used a regular data plan provided by the phone's carrier.

Overall, the creation of this system followed a typical software development pattern. To summarize, in our first live pretest we manually simulated how the proposed software would work in practice. Next, we built the software and tested it using AOL's instant messaging network. A month after the initial simulated pretest, we did another pretest using our first service provider and discovered its reliability problems. We spent several months researching alternative service providers and then building our own homemade text messaging setup that used our own mobile phone. We tested this configuration of hardware and software and found that it was quite reliable. A week later, we went live and started collecting data from respondents.

All-in-all, the main take-away message is that researching available tools is important, but one should not compromise core needs of the project just to cut down on some initial up-front investment in tool development. Moreover, continuous testing of the instruments is essential for addressing the various issues that arise during such an undertaking.

Developing and refining the coding scheme

As noted earlier, the overarching substantive goal of our study was to get a better understanding of how digital media intersperse adolescents' everyday lives. We collected diary data with the intent of seeing how often college students in our sample use various digital devices and in what types of communication processes they engage during the course of a day. That is, we were interested in seeing the extent to which they spend time watching television, using the Web, interacting with others face-to-face, and so on. Additionally, we

[4] Since this could end up being an imposition on the person whose phone is thereby taken up all day for the study, future studies may want to opt for purchasing a separate phone and data plan for the project.

wanted to learn about this in the context of their other activities. Ultimately, this meant creating a coding scheme that would account for anything anyone might do. Obviously this is a daunting task with unlimited options. We made this task manageable by deciding on the categories of information most relevant for our purposes and we classified responses by type. We will not get into the specifics of the coding scheme here, but we want to say a few words about how we developed it and refined it during the study.

Text message responses had to be coded by more than one person to establish inter-coder reliability. We trained two undergraduate research assistants for this task in addition to doing a bit of preliminary coding ourselves. Similarly to when we were developing the interface, we asked research team members to give us feedback about both the tool and any difficulties or ambiguities posed by the messages they were coding. We decided to rely mainly on email for communication and a flurry of messages soon flooded our mailboxes. Prompt responses were important so the coders could proceed with their job. We realized that several types of issues were cropping up with some regularity so we decided to take a closer look at the entire coding scheme together.

We met using Skype and made significant progress. While email can be extremely helpful, it is hard to replace the efficiency that can be achieved in one or two hours spent on questions of this sort face-to-face. This holds even when not all participants are physically co-present. Soon after our initial meeting, we held another meeting and came close to finalizing our scheme with only a few tweaks left. A few more minor adjustments surfaced in the following days, but soon we were able to finalize the coding scheme.

After these alterations, the research assistants started over with their coding. We made it clear to them that this was in part a methodological project so they should not feel that energy spent on coding that was now discarded had been wasted effort. It was important for their morale that they understood that the feedback they had given us was an integral part of the project, that it was an important part of their job, and that their input was taken very seriously and was much appreciated. In fact, it was essential throughout this exercise to let research team members know that we took their comments very seriously and encouraged their contribution.

Despite every attempt to make responses systematic, the reality of data collection is never as clean and straightforward as one envisions up front. While the majority of the responses we received were sent shortly after the participants got our requests for

information, some came in considerably later. This was mainly due to issues we discussed above about disruptions in people's service (whether due to technical unavailability or a conscious effort to disconnect in some situations). Consequently, some responses came in after we had already sent subsequent requests. These cases were usually easy to note since they entailed receiving a quick succession of responses from a participant separated by minutes only as opposed to the standard hour or so difference between messages. These responses were typically sent as a batch when the respondent finally got around to responding to requests missed. We interpreted responses in order (unless the participant specified a time stamp in the message that led us to believe we had not received the messages in order) so overall this did not pose a major challenge.

Another issue we had not anticipated and one that was not trivial to handle concerned responses that referenced information communicated to us earlier in the day. That is, on occasion we would receive a message that may have simply said something like "same as before" or "still at work" without further elaboration. In such cases, we may have already possessed additional information about the setting, but we had to decide how to code the entry. We decided to add fields to the coding scheme signaling whether different types of information had been included in the message itself, even if the information was known to us but not made explicit in the short response.

All-in-all, this was a very detailed and valuable exercise. Getting the coding scheme right is crucial to a study's success. In particular, it is important not to lose information about the data at this stage. That is, it may be that later in the project we decide to get rid of certain nuances in the data set by collapsing various categories or values. Nonetheless, not knowing all details of potential analyses ahead of time, it is best to hold on to as much nuance about the data as we have at our disposal. Collapsing and aggregating material is always a possibility later whereas any information lost during coding remains lost to all subsequent investigations (unless one goes back to the raw data, which is not realistic in most cases given the effort involved).

Building and refining the coding interface

With the data collection components in place and having an idea of what information we wanted to extract from the collected data in a systematic way, we began developing a tool for coding and annotating the responses we would receive from participants. We wanted to be able to create a flexible and user-friendly interface. Since we

would have several people working on the coding – on occasion concurrently – we also wanted a tool that was accessible from within a Web browser and could be used by more than one person at a time. This remote accessibility allowed coding to take place from different locations.

To provide a rich interface that avoided the pitfalls of cross-browser incompatibilities, we used Adobe's Flash as our platform. We created a tool that directly imports the data collected by the scheduler and builds an interface that reflects the desired coding scheme. The coding scheme is saved as an XML file that the tool interprets to construct a suitable interface. Components of this interface can be as simple as a checkbox or as complex as a tree view that allows multiple selections. A paging mechanism allows the interface to represent coding schemes of arbitrary length, thereby not imposing a technical limitation on the scheme authors. Again, by using XML, we were able to test and development the scheme rapidly and iteratively. Had we used an external database or other file format, this would have delayed our development by introducing additional installation and integration requirements. Our format allowed us to add new fields and options by simply updating the file in a text editor. This proved useful when coders recognized that the scheme was missing crucial items that needed to be included.

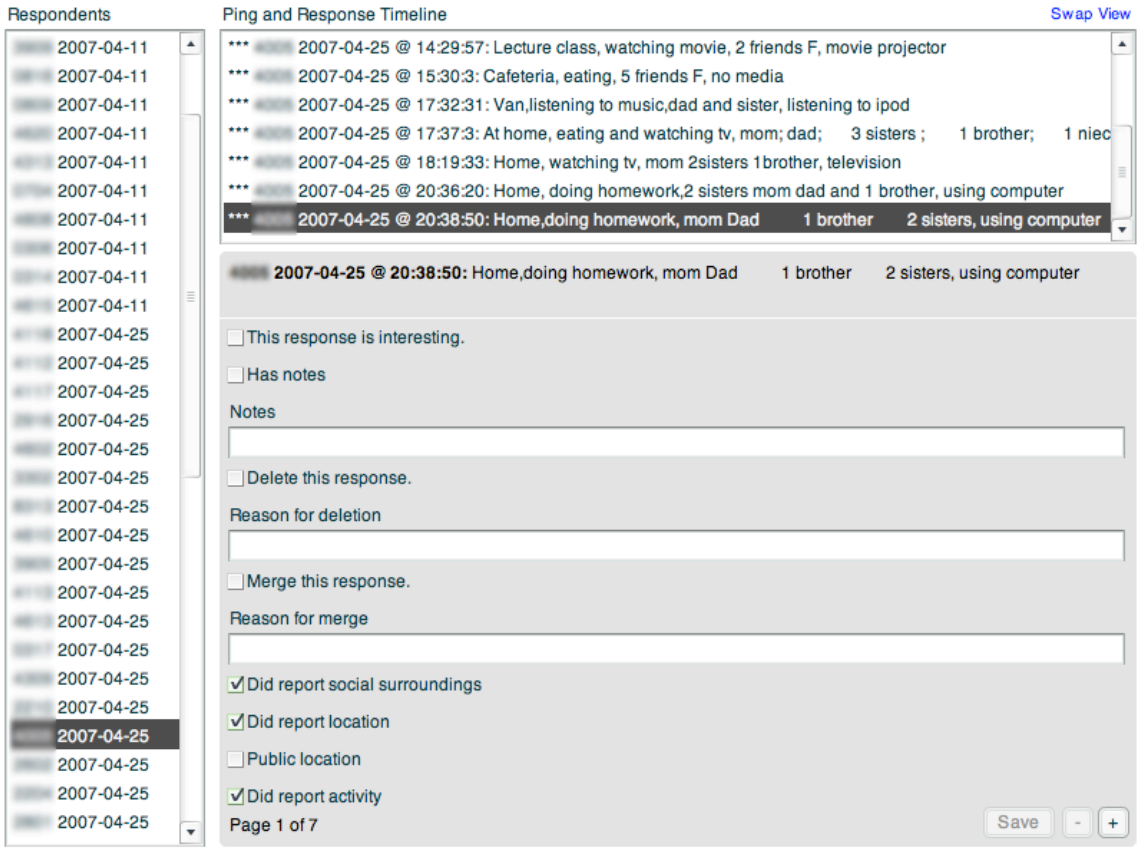

Figure 1: The Web-based coding interface

Since we previously obtained the Flash software for a prior project, we incurred no costs when creating this interface. However, we should note that unlike the Java tools used to build the scheduler, the Flash tools do cost several hundred dollars and this should be taken into consideration if they are unavailable. Overall, it took us about two weeks of part time work to create the initial version of the coding tool.

Once we finished development of the tool, we tested it to learn how well it met our needs. We trained an undergraduate research assistant and using data collected from our pretests, we started coding responses. There are always unanticipated aspects of coding – both at the substantive and technical level – that are impossible to predict without putting a tool into action. We made it clear to the research assistant that this was a testing period and so all comments about the interface would be welcomed. She was also encouraged to ask questions about any ambiguous text message responses that she was not sure how to handle. It is important to have this kind of open communication among team members so that no individual is left making important decision in isolation. Input started coming in soon after we started this exercise and we promptly made changes to the tool to address the various

concerns having to do with both the usability of the tool and the coding scheme (whose compilation we described in detail in the previous section). We received another round of feedback at a later stage in the project when we moved on to coding responses from the actual study. Similarly, we responded continuously to all suggestions. That round led to some more elaborate changes to the system, for example, the addition of a tree-view widget to facilitate navigation between coding options.

All-in-all, having our own interface to code the responses worked out well. We were fortunate to have a team member who had the necessary expertise to implement the interface. Something similar could likely be achieved by hiring someone part time. In addition to the added financial cost, the downside of such a solution, of course, is that the person would not be available as readily and promptly as a member of the research project. Based on our experiences, it is worth having such a customized platform. With the refinements suggested by the research assistants who were using the interface the most, the tool helped speed up coding considerably.

Data Collection

Sampling

Since this study was part of a larger data collection effort, we did not face the task of creating a sampling frame. The sampling frame was the group of students who had been recruited into the observational and interview component of a yet larger study examining adolescents' digital media uses, skills and participation. The overarching study was based on very rigorous sampling methodology so we knew that we were working with a random sample of students from a well-defined population. Namely, the study was based on students enrolled in the one required course of an urban public research university. That study had achieved a high response rate with the follow-up segment also successful in recruiting students into the observational component of the project. Therefore, we knew we would get a wide representation of students in this portion of the study.

Working with the researchers who were conducting the interviews with participants, we recruited students into the text-messaging study at the end of the observational session by asking whether they would be willing to take part in an additional study. If they agreed, we asked them to sign a separate consent form for this study and gave them compensation for their anticipated text-messaging fees. Arranging these logistics at this stage of the

process was important given that the whole point of this methodology was to avoid the need for physical co-presence with participants for the actual data collection.

Compensation

Needless to say, paying people up front for their participation in a study is tricky since it raises concerns about respondents running off with the money without meeting their end of the agreement. We were nervous about this and addressed it by providing an incentive for participation that students would get after the data collection. We gave respondents $10 cash up front for subsidizing their text-messaging fees associated with the study. Additionally, we committed to sending them a $15 gift certificate at the end of the data collection. We also conducted a drawing for an iPod.

Integrating participant information

Since we wanted to study the participants throughout their waking day, we needed to know when people would be awake so that we could tailor the schedule of messages to fit their day. This is information one could collect about participants when they agree to be in the study, but we had neglected to do so. Contacting respondents for this information gave us the opportunity to remind people of the study. We had participants' cell phone numbers from when we had recruited them into the study so we were able to contact them both via voice and through text messaging, in addition to email, if necessary.

We used a simple spreadsheet for keeping track of all relevant information about respondents, which concerned participants' mobile numbers, availability during the day of the study, and email address.[5] The day before the study, we filled in any missing information about when respondents would be awake with our default values for availability and we translated the information into a short Ruby script that scheduled the messages on the scheduler Web application. These scripts rarely exceeded seventy lines of code. The day before the study was also dedicated to setting up the system and conducting some preliminary tests to confirm that there were no problems with the configuration. The evening before the study, we scheduled a handful of messages to be sent to our own mobile phones. This allowed us to verify that we received the scheduled messages and that the

[5] For confidentiality purposes, any such information was always stripped of identifying information so we only had ID numbers and cell phone numbers without any names. These documents were kept in password-protected directories on university computers to which only research team members directly involved with this project had access.

software received and cataloged our responses. After we were confident that the system was sound, we ran the study script to schedule the messages for the upcoming data collection.

Reminders to participants

Several days before the day of the study (Wednesdays for these four waves), we e-mailed the people on our list for this information. In the case of our first wave, we started contacting people on Monday for Wednesday participation. This proved to pose a challenge since students either were not reading their mail regularly or were not responding in a timely manner for other reasons. During later waves, we wanted to be more certain to leave enough response time so we started making email contact on Fridays. While most of these email confirmations verified the mobile numbers we had on hand, we did find a few errors where a mobile number had been miscommunicated earlier so it was helpful to double check such a crucial detail. In addition to the request for information about waking hours, these email messages also included reminders about the study as a whole, the goals of the project, and information about how to participate.

Even with the added lead time, not all participants responded to our requests in a timely manner and in the days leading up to the study we resent the message until we heard from them. It was also useful for the lead investigator to email the students to remind them that they had already been paid for their participation and so while they were certainly free to decline participation – an important caveat in compliance with human subjects protection guidelines – they would have to return the money to us if they did not take part. (Of course, there was no way for us to pursue the money if students decided to back out and not send the $10, but it was worth a mention.) Keeping a polite tone and explicitly acknowledging that the study was voluntary was important throughout this communication. If we did not obtain information about hours from a respondent then we scheduled the person's participation in the study using a default time window that stretched from ten o'clock in the morning until ten o'clock at night. This was extremely rare, however, and we only had to resort to calling people up to check on participation in a few cases.

Finally, to test the system and get the respondents into the mode of communicating with us through text messaging, the night before the study we sent a reminder to respondents' phones noting that the study would begin the next morning and suggesting that they add the message's sender to their address book. The challenge of writing this reminder message – and all other messages we sent – was that such messages should not be more than

160 characters to comply with restrictions that some phone companies put on the length of text messages. We ended up using the following 133-character text as the reminder message the evening before the study: "Tomorrow we will be conducting the SMS study you signed up for. Please respond to all messages you receive from this number tomorrow." We avoided using shorthand messages in case any of our participants lacked familiarity with them.

Going live

Using text messaging with the system we built, we collected diary data from 60 respondents in four waves over the course of three months. Because this project supplemented another one and relied on it for recruitment into this study, our timeline was dependent on the other process. Were it not for this constraint, the entire study could have been run much more quickly. We conducted this study in several waves, because we wanted to make sure that data collection occurred close to the time when respondents were recruited into this study so that they would not forget about their participation.

The first data collection took place in the middle of April (15 participants) with the second occurring two weeks later (20 participants). Our third wave of data collection had to be sensitive to the academic calendar of the institution at which our respondents were studying. We had to wait a month between the second and third waves to accommodate the final exam schedule. Consequently, the third wave occurred during the initial weeks of the participants' summer vacation, in late May. While we managed to engage all of the people who had signed up for our first two waves, our third attempt proved less successful. Of the twenty-six people who signed up to participate, five did not respond to our text-messaging requests. We recontacted the absent participants and convinced all of them to enroll in a fourth "make up" study two weeks later. Four of those five responded to the messages in this last wave.

Participants received several different types of messages during the course of the data collection day. A few minutes before our first request for information, we sent a message alerting the participant that the study was about to begin. The 136-character message read as follows: "Good morning. Thank you for agreeing to participate in our text messaging study today. You will receive our first request for info soon." After the morning reminder, we sent messages hourly requesting that the participants respond with their location, activity, social context, and any media in use. Again, we faced the challenge of

fitting the instructions into a very short message, this one 147 characters long: "Please reply with your location, current activity, people you are with (number, your relationship to them, gender) and any media you are using now." We sent these message fifteen minutes after the hour so that we avoided capturing any non-typical activities that may be associated with the beginning of an hour (start of a work shift, a class, or a meeting, to name a few).

After the last request for information had been transmitted, we sent a final message a few minutes later informing participants that the study was over, thanking them for their participation, and providing contact information in case they had any questions. ("The SMS study is finished for the day. Thank you for your participation. Questions? Call: xxx-xxx-xxxx.") We also included ourselves in the list of participants and we received all of the same messages as the participants. This was very useful throughout to identify any problems in the transmission process. Fortunately, we encountered none.

While we encouraged participants to respond to the messages as soon as possible, we realized that immediate responses would often be infeasible. Students may be in class or a meeting and unable to respond. Alternatively, it was possible that they would receive our messages while out of range. We instructed them to respond at the next nearest time they were able to do so with information on what they had been doing at the time they had received the message. To eliminate any confusion about the time a message was sent and because not all cell phones include an automatic time stamp on text messages, we included this information as the first few characters of each text message we sent out to respondents.

The process of running the study was largely automatic and only required one team member to contact the participants, collect their information, schedule the study, and monitor the study for any unforeseen problems. The bulk of the work was concentrated in the days before the study, with the majority of the effort focused on establishing contact with the participants. A moderate amount of effort was involved with writing the scheduling script, but this took less than an hour for the base script and less than twenty minutes for customization with any given wave of data collection. It was important for one of us to be present with the system during the day of the study, but monitoring the progress was a background task that only required attention every hour or so. Shutting down the study the next morning required that we archive the collected data and shut down the software. This typically took less than a half hour. Overall, all four waves of our data collection went

smoothly, something we attribute to our extensive testing and tweaking leading up to data collection.

Data processing

After shutting down the study, we moved the collected data to a university-based shared digital space for later use. In preparation for coding the collected data, we first combined the responses into a single file and cleaned the data set by getting rid of text messages that were not substantive in nature. (The participants often sent simple "ok" messages to our reminders about the study. We discarded these so as not to clutter our coding process later.) We used the resulting aggregated master file as the data set for coders.

Lessons Learned

In many ways the actual data collection is the most essential component of such a research project so how is it that the section (see Going Live above) devoted to it in this chapter is one of the shortest? This highlights the importance of careful preparation leading up to data collection. Respondents' time and attention is at a premium. Glitches occurring at that stage of the project can be fatal to a study. Therefore, it is imperative that researchers put much care into all phases of the undertaking leading up to the crucial moments where participants are directly involved in the data collection.

We learned several important lessons during this project. Communication with respondents may be easier to achieve using a combination of media (email, voice, text-messaging), but one-on-one attention remains important regardless of the particular means of contact. That is, although we relied on automated template messages – with personalized greetings – to establish contact, it became clear that respondents often required additional information whose delivery would be hard to automate. This has implications for the scalability of the project. If one were to try such a study with thousands of respondents, it would be essential to devote resources to one-on-one contact with participants given the number of issues that tend to come up and that require a resolution before data collection can proceed (e.g., clarification on both ends of logistical details about the study including means of subsidizing text-messaging costs, timing of message exchange, costs of messaging, and the timeline for reimbursement).

In a technical sense, we confirmed the fact that open source and open standards are important tools for developing technical solutions to research problems quickly and cheaply.

With the exception of the Flash coding interface, we built our entire system using free software available online. We resorted to Flash since it had better compatibility between browsers than the alternatives. If Flash were not available to us, we may have investigated more seriously the use of dynamic AJAX interfaces instead. We also confirmed that creating open and extensible architectures from the beginning of the project is very important. This allowed us to prototype and test the system with a readily available free network while we investigated more robust commercial alternatives for the actual study. We were able to adapt our system for the text-messaging services we found, but we were not locked in and this allowed us the flexibility ultimately to create our own substitute service. An extensible architecture within the coding tool allowed us to extend our interface with a tree-view later in the process when we found that a simple list was not efficient from the coders' perspective.

Engineered extensibility is not only limited to the software and source code. By adopting a format that facilitated an easily customizable and extensible coding scheme, we have been more nimble and responsive in the development of our scheme. This proved useful when we identified information that we were not previously capturing or options that we initially overlooked. Our coding scheme benefited in the same ways as software development when using a tight iterative cycle. Our scheme is more complete and was more responsive to the issues that our coders identified.

We also discovered that while the mobile text messaging and instant messaging networks appear quite similar, this is not the case. The text messaging network is quite proprietary and requires more capital and work to establish a presence. If we wanted to create a presence on the network with a minimum number of middlemen and resellers, we would need to spend tens of thousands of dollars (and several months) to obtain a short code. Since this was beyond our means, we were forced to deal with resellers with their own short codes. These services are still expensive and the resellers focus more on the lucrative marketing projects than the typical academic study. However, in the end, we discovered that we could still participate on this network through the creative use of a single mobile phone hacked together with some open source tools.

We found that the Thumb Generation is comfortable with participating in this type of study and that our greatest difficulty was not dealing with issues such as privacy concerns or text messaging costs, but instead dealing with the limitations of the technical format and

ingrained patterns of use that stressed the use of short messages. The respondents often provided all of the information we asked for (or numerous times even more), but they rarely repeated it twice, instead relying upon messages like "same as before" to show that they were in the same situation. We have dealt with this as it appears in our coding, but a future revision of this project may wish to address this in a more fundamental manner by including technical tools that recognize and respond to this situation in a more automated manner when possible.

A caveat must be made at this point about the generalizability of this study to other projects when it comes to the content of the information we collected. It is important to remember that we piggybacked on a larger project in which significant amounts of information had already been collected about our respondents. Therefore, we were not dependent on collecting baseline demographic data, to name one example, about participants using this method. Studies most likely to benefit from our experiences are ones that also use another methodology to collect some background material about respondents and then use text-messaging for follow-up data collection.

Collaborative work can have both very rewarding and very frustrating components. We managed to avoid the latter thanks to a deep commitment to the project on behalf of team members, frequent and respectful communication, and explicit idea exchange. We considered each other's feedback seriously and when not on the same page initially, we explained, patiently, the reasoning behind our positions in a detailed manner. Being comfortable with asking questions of others on the project was very important, especially given our different disciplinary backgrounds. Pretesting various components of the study allowed us to address unanticipated challenges in a timely manner. Because various steps of the project are so dependent on each other (e.g., the coding interface is directly linked to the coding scheme), leaving the revision phase to the last minute would have left us with much to do and would have delayed the process as a whole. Not collecting initial data from our own trusted group of team members would have also jeopardized the quality of data we collected from study participants. Actively seeking input from our research group and research assistants was also essential to being able to make the types of quick improvements to our coding scheme and interface that allowed continuous progress.

Finally, it is worth noting that being involved with every step of the process is important for having realistic expectations of what work is involved in a study from building

the technical aspects of the system, to what data are realistic to collect and how they should be handled. We both took part in simulations of data collection, testing of the coding interface, compilation of the coding scheme, and communication with respondents. (Of course, we did the latter in a coordinated manner that presented a unified front to participants.) While we were certainly not equally involved with each aspect of the project (e.g., Karr gets credit for the programming work that went into building the technical systems), we both had a realistic idea of what we were asking of each other, what we were asking of our research assistants, and most importantly, of what we were asking of our respondents. There is no substitute for such direct involvement and it adds significantly not only to the final research product, but also to the new skills and know-how the researcher is able to take away from such an experience.

References

Barrett Feldman, Lis, and Daniel J. Barrett. 2001. "An Introduction to Computerized Experience Sampling in Psychology." *Social Science Computer Review* 19:175-185.

Carter, S., J. Mankoff, and J. Heer. 2007. "Momento: support for situated ubicomp experimentation. ." in *SIGCHI Conference on Human Factors in Computing Systems*. San Jose, CA.

Christensen, T. C., and L. Feldman Barret. 2003. "A Practical Guide To Experience-Sampling Procedures." *Journal of Happiness Studies* 4:53-78.

Pentland, Wendy E., M. Powell Lawton, Andrew S. Harvey, and Mary Ann McColl. 1999. *Time Use Research in the Social Sciences*. New York: Kluwer Academic/Plenum Publishers.