WAT R U DOIN?

Studying the Thumb Generation Using Text Messaging

ESZTER HARGITTAI AND CHRIS KARR

A relatively recent and sudden change in the landscape of American youth can be seen in the growing numbers of people walking around with their heads looking down and their fingers moving vigorously on a gadget. Cell phones have spread widely, and their use for communication through text messaging has taken off considerably. Is there a way for social scientists to benefit from the proliferation of this technology? Can such short messages help us understand human behavior better?

We were in the midst of an unrelated study when we suddenly realized that we could piggyback on it to supplement survey questions with time-diary data collected using text messaging from college students about their everyday activities. The method of data collection that relies on calling respondents has suffered from declining response rates for years, while more traditional time-diary data collection means (e.g., journaling or beeper studies) have posed their own set of challenges regarding logistics and data quality. We were curious to know if text messaging, a new method of communication already present in many people's everyday lives, might allow us to improve on existing methods.

With this in mind, we decided to add a component to a larger study that was already under way with questions that we believed were worth pursuing even if, in the end, our time-diary data collection ended up being exploratory at best. Overall, we had both methodological and substantive reasons to pursue this new 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 that we had seized this opportunity, and although it took more time—as it always does!—than we had anticipated, we achieved interesting and unique results that were well worth the investment.

At the time we embarked on this research, one of us (Hargittai) was in the midst of working on a two-year project that involved studying adolescents' Internet uses, skills, and participation by means of surveys and in-person observations. The study also had a longitudinal component whereby some participants would be randomly assigned into a training program and then observed again at a later point in time—along with those who had not received training—to test whether the intervention had made a difference in students' online know-how. These parameters meant that some people were already going to be approached for participation more than once. More important, the data that we had begun collecting on respondents could easily be merged with the additional information we were hoping to collect about them through 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 (e.g., Larson and Csikszentmihalyi; 1983, Robinson 1977). However, given students' busy lives and the difficulty in convincing people to participate in recurring studies, the challenge remained: How do we collect diary data from 18- and 19-year-olds who are physically hard to pin down amid 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 either calling people in their networks or texting them using the dial pad of their phones. Data that one of us (Hargittai) had collected a year earlier about a similar group suggested that most students in the population of the larger study owned cell phones, and many used text messaging. Accordingly, our belief that texting was a popular activity was not simply based on unsubstantiated assumptions but rather, available data. In fact, a look at the survey re-

sponses of the current study's sample made it clear that over 98 percent owned a cell phone and that over 90 percent of those cell phone owners used the device for text messaging. This gave credence to the idea that collecting diary data through the relatively unobtrusive medium of text messaging might yield helpful information. The method would not require researchers and respondents to be physically copresent, and it would draw on an activity in which students are already engaged during their everyday lives using a device they already own and operate.

In what follows, we will describe the various research tasks associated with this project. We will say a few words about the long-distance nature of this collaboration. Next, we offer a full time-line of the research project in order to give 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, as 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 go on to 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 the kinds of challenges that may be encountered when trying to scale up from our experiences.

LONG-DISTANCE INTERDISCIPLINARY COLLABORATION

While somewhat tangential to the study, one important point we want to get across is that long-distance collaboration is very much feasible in this day and age and that researchers should not be deterred from pursuing a joint project simply because they are not physically copresent. The lead on the overall study was a junior faculty member at the time (Hargittai) on leave 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 PhD 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 coauthored 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 did pose some challenges especially when attempting to communicate certain ideas through our different disciplinary terminologies (see also Sandvig's piece in this volume). However, frequent communication—mainly using e-mail—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 one of us was on leave two thousand miles away from the other and from 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 instructive examples of collaboration, but long-distance communication and coordination among multiple people and project components. We want to acknowledge the important role that free online services, such as the videoconferencing tool Skype, play in making such undertakings possible.

Finally, while only two names appear on the byline 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 an acknowledgments section, 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 for the larger undertaking 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 latter point later in the chapter.

TIME LINE

Before launching into a detailed description of how we approached the various components 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 want to present

an overall time line of the project (table 1). Our goal here 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 that rarely ever see the light of day in publications.

TABLE 1. Time Line of the Project

	Nov	Dec	Jan	Feb	Mar	Apr	May	Jun	Jul
Project idea, first e-mail 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									
First internal pretest on research group									
Constructing and fine-tuning our coding scheme									
Training coders									
Fine-tuning our coding interface									
Coding of pretest data (to test coding interface and scheme)									
Second internal pretest on research group (expanded group)									
Recruitment of respondents									
First data collection (15 participants)									
Second data collection (20 participants)									
Third data collection (21 participants)									
Fourth 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 when the researcher almost inevitably assumes that 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 automated text-message requests to which respondents could reply. As the section on setting up our system attests, while not impossibly difficult, this 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 are 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, 27). However, the difference here is that we had a particular focus on communication processes and digital media use. Moreover, our respondents were constrained by the limitations of the medium. Is it realistic to expect such detailed information from respondents in 140 to 160 characters, which is the message limit imposed by service 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 students' lives. However, as our notes will demonstrate, gathering this type of information from participants through this re-

stricted medium requires communicating detailed instructions to them ahead of time, and that raises its own set of 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, participants' mobile phones. Our "pinging" system—as we called it, in a nod to the practice of sending short messages to networked machines to assess their availability—would have to meet two main criteria. First, we had to automate the process of sending out requests for participant responses. That is, we needed the ability to schedule the requests ahead of time and implement 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 wanted to run the study—as we did—multiple times on larger groups 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 that were targeted to studies similar to our own in the hopes that we might be able to use them for our purposes. In the end, this process did not reveal any systems that we could adopt for our study though it did help us clarify the needs of our project. We include this part of the process here because it is precisely the kind of detailed description that is never included in project write-ups but nevertheless 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 tool kit 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 ultimately a poor fit for our needs. The fact that it originated in the human-computer interaction field meant that it was primarily a tool for simulating interactions on mobile devices rather than 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 (Feldman Barrett and Barrett 2001), a tool kit 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 its reliance on devices that participants already owned, requiring specific pieces of equipment or programs was a significant deterrent. We did not want to be hindered by requirements of particular hardware or software specifications beyond what would be readily available to anyone who has a text-messaging subscription on a regular cell phone plan. Providing study respondents with such devices would be cost prohibitive, and introducing a new device would recreate many of the problems associated with traditional paper-based or beeper diary studies (see, e.g., Pentland et al. 1999; Christensen and Feldman Barrett 2003). For example, 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 (for complexities involved in such undertakings, see Adam, Doane, and Mendelsohn in this volume).

We also evaluated several commercial services for sending and receiving messages, but we found that the services were either too limited, too unreliable, or too expensive (or some combination of these three) for our needs. We investigated a number of other providers, which supply services that avoid problems such as spam filtering, but these services were both too expensive and a poor fit for our needs, having been designed primarily 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 re-

lationship than was compatible with our study's time line. Finally, we were mindful throughout this initial process of the need to avoid becoming locked into 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 of the system.

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 quickly 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 larger project, and given that we have a collaborative atmosphere in the lab, most research group members willingly participated, giving us helpful feedback. We participated in this pretest ourselves as we believed it important to acquire firsthand experience with the method from the perspective of a respondent.

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 140–160 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 for feedback about what they were doing on an hourly basis. We also solicited input from the pretesters 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.

We 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 predefined times. We created a simple Web-based interface that allowed the manual scheduling of requests, but we also provided a remote application programming interface to be used by external scripts to batch-schedule complete studies.[1] 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 stand-alone 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—that is, 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 since the way the provider sent our messages activated the spam defenses on the mobile phone networks.

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.[2] 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 very reliable. A week later, we went live and started collecting data from respondents.

All in all, the main takeaway 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 are integrated into adolescents' everyday lives. We collected diary data with the goal of seeing how often college students in our sample use various digital devices and what types of communication processes they engage in during the course of a day. That is, we were interested in seeing how much time they spend watching television, using the Web, interacting with others face to face, and so on, and how reports about these activities through the text-messaging medium compare to participants' survey responses about their media uses. In addition, we wanted to learn about digital media uses and other communication practices 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 do 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 intercoder reliability. We trained two undergraduate research assistants for this task in addition to doing a bit of preliminary coding ourselves. As when we were developing the messaging system, here again 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 (see the next section for details about developing the technical tool we used for coding). We decided to rely mainly on e-mail 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 coding scheme as a whole.

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

After these alterations, the research assistants began the coding process again. We made it clear to them that this was in part a methodological project so they should not feel that the time and effort spent on coding that was now discarded had been wasted. It was important for their morale, we

believed, that they understood how 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, had been incorporated into the project, 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 contributions.

Despite every attempt to make study participants' 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 disruptions in people's service (whether due to technical unavailability or a conscious effort to disconnect from the network in some situations, e.g., during class time). 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 participant responses that were separated only by minutes 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 by participants in earlier messages during the day. On occasion we would receive a message that may simply state "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 as it did not itself contain additional information.[3] We decided to add fields to the coding scheme signaling whether the four main 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. 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 detail about the data as we have at our disposal at the initial coding stage. 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 FINE-TUNING THE CODING INTERFACE

With the data collection components in place and an idea of what information we wanted to extract from the collected data systematically, we began developing a tool for coding and annotating the responses we would receive from participants (fig. 1). 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 develop 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 team members recognized that the scheme was missing crucial items that needed to be included.

Since we had 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 money is 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

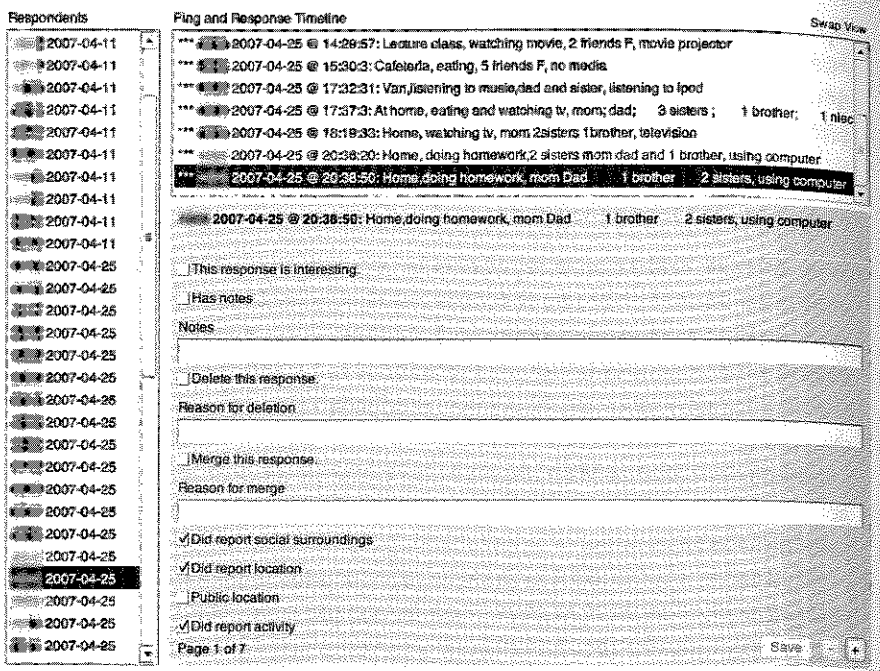

Fig. 1. The Web-based coding interface

well it met our needs. As described in the previous section, we trained undergraduate research assistants, and using data collected from our pretests, we started coding responses. There are always aspects of coding—both at the substantive and technical levels—that are impossible to predict without putting a tool into action. Input from the research assistants started coming in soon after we started this exercise, and we promptly made changes to the interface 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. Again, we responded promptly to all suggestions. That round led to some more elaborate changes to the system such as, 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 (Karr) 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

Getting IRB Approval

Before we could proceed with going in the field, we had to obtain permission to do so from our Institutional Review Board for Human Subjects (IRB). Since we were piggybacking on an existing project that had already obtained IRB approval, we submitted a revision request to the already approved study. It turns out that gaining approval for a revision to an existing project can be much quicker than starting up a new study. This makes sense since the board will have already reviewed the overall framework and logistics of the project. With experience, we also knew what specifics (all of them!) the IRB would want from us, and so we were sure to include everything in detail in order to avoid having to revise the initial submission. This worked out well, and within a few weeks we were approved to proceed with 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 for the observational and interview component of the larger study examining adolescents' digital media uses, skills, and participation. The overarching study was based on very rigorous sampling methodology, and so we knew that we were working with a random sample of students from a well-defined population. More specifically, the study was based on students enrolled in the one required course at the University of Illinois, Chicago. That study had achieved a high response rate of 82 percent, and the follow-up segment was similarly successful at 53

percent in recruiting students into the observational component of the project. Therefore, we knew that our participants would be a diverse group.

Working with the research assistants who were conducting the interviews with the participants in the observational study, we recruited students into the text-messaging project by using the end of the observational session as an opportunity to ask whether they would be willing to take part in this additional study. If they agreed, we asked them to sign a separate consent form for the texting project and gave them compensation for their anticipated 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 copresence 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. In addition, we committed to sending them a $15 gift certificate at the end of the data collection. We also conducted a drawing for an iPod where each text message sent to us in response to our requests added to the likelihood of a person being picked the winner. Our hope was that this would provide an added incentive for participation. Overall, most students readily participated, and only in a few cases did we need to send a few reminder messages (see below for more details on this point).

Integrating Participant Information

Since we wanted to study the participants throughout their waking hours, we needed to know when people would not be asleep so that we could tailor the schedule of messages accordingly. 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' e-mail addresses and cell phone numbers from our original recruiting, and so we were able to contact them using both voice and text messages, in addition to e-mail, if necessary.

We used a simple spreadsheet for keeping track of all relevant information about respondents, which concerned participants' e-mail addresses, mobile numbers, and availability during the day of the study.[4] The day before the study, we filled in any missing information about when respondents would be awake with our default values for availability (10:00 a.m.–10:00 p.m.), and we translated the information into a short Ruby script that scheduled the messages on the scheduler Web application. These scripts rarely exceeded 70 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 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

Given the time that elapsed between recruitment of respondents into the study and the date of our data collection, it was important to remind participants of the project a few days before going live with data collection. We also wanted to verify that the cell phone numbers we had been given were in order and that we had people's awake hours to know when to send them our automated pings. Accordingly, we had to leave enough response time between the time of our notification and the study to hear back from students. After realizing from our first run that contacting students on Monday for a Wednesday study may be cutting it close, we started making e-mail contact on Friday for a Wednesday study.

While most of the e-mail confirmations we received verified the mobile numbers we had on hand, we did find a few errors where a mobile number had been miscommunicated, and so it was helpful to double-check these crucial details. In addition to the request for information about waking hours, these e-mail messages 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 to those who had not gotten back to us until we received

a response. It was also useful for the lead investigator to e-mail 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 back 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 our 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, we sent a reminder to respondents' phones the night before the study 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 had to comprise less than 140–160 characters in order 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 were unfamiliar with them.

Going Live

Using the system that we built for text messaging, we collected diary data from 60 respondents in four stages over the course of three months. Because this project supplemented another one and relied on it for recruiting purposes, our time line was dependent on the logistics of the larger project. If we had not had this constraint, the entire study could have been run much more quickly. We conducted the study in several phases, because we wanted to make sure that data collection occurred close to the time when respondents were recruited into the study and thus less likely to forget about their participation.

While we managed to engage all of the people who had signed up for our first two groups, our third attempt proved less successful. Of the 26 people who signed up to participate, 5 did not respond to our text-messaging requests. We recontacted the absent participants and all of them enrolled in a fourth makeup study two weeks later. In the end, 4 of them responded to the messages in this last round of data collection.

Participants received messages from us 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 messages fifteen minutes after the hour so that we avoided capturing any nontypical activities that may be associated with the top of the 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 103-character 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 in order to ensure that we received all of the same messages as the participants. The goal was to identify any problems in the transmission process immediately. 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 might be in class or a meeting and unable to respond. It was also 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 in 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 case 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 half an hour. Overall, all four instances 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 network 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

Since the actual data collection is in many ways the most essential component of such a research project, it is worth explaining why the section (see Going Live above) devoted to it in this chapter is one of the shortest. The lengthy description of doing the groundwork 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 and data collection begins.

We learned several important lessons during this project. Communication with respondents may be easier to achieve using a combination of media (e-mail, 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 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 time line 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 modifiable coding scheme, we were more agile and responsive in the development of our scheme than would have been possible otherwise. This proved useful when we identified information that we were not previously capturing or options that we had initially overlooked. Our coding scheme benefited in the same way 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 to be quite similar, that is not in fact 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 in 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 difficulties were not caused by privacy concerns or text-messaging costs but the logistics of finding suitable software and the day-to-day logistics of setting up the project. Drawing on our lessons learned, however, we were able to run a second iteration of the project a year later much more efficiently. In this second study, we had 75 participants and collected data about each of them during three days (two weekdays and a Saturday). Overall, that study greatly benefited from the care we put into developing the details of the project in the first round.

A caveat must be made at this point about the generalizability of this study to other projects, particularly as it relates 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 information 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 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 important, 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.

NOTES

The authors thank Viktor Domazet, Alex Knell, and the members of the Web Use Project research team of 2006–7 for their helpful input. They also acknowledge the assistance of Elizabeth Anderson, Waleeta Canon, and Gina Walejko. The authors are grateful 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 also thanks 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.

1. 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 in the future 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.

2. 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.

3. Since we coded messages manually, it was relatively easy to identify such cases. In the future, if machine coding was implemented, messages of this sort may pose a special challenge.

4. 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.

REFERENCES

Carter, S., J. Mankoff, and J. Heer. 2007. *Momento: Support for Situated Ubicomp Experimentation.* SIGCHI Conference on Human Factors in Computing Systems, San Jose, CA.

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

Feldman Barrett, L., and D. J. Barrett. 2001. An Introduction to Computerized Experience Sampling in Psychology. *Social Science Computer Review* 19:175–85.

Larson, R., and M. Csikszentmihalyi. 1983. "The Experience Sampling Method." *New Directions for Methodology of Social and Behavioral Science* 15:41–56.

Pentland, W. E., A. S. Harvey, M. P. Lawton, and M. A. McColl. 1999. *Time Use Research in the Social Sciences.* New York: Kluwer Academic/Plenum Publishers.

Robinson, J. P. 1977. *How Americans Use Time: A Social-Psychological Analysis of Everyday Behavior.* New York. Praeger Publishers.

GIVING MEGA ATTENTION TO MACRO RESEARCH

The Rewards and Challenges of Quantitative Cross-National Data Collection and Analysis

NINA BANDELJ

Comparison is central to any social scientific inquiry. As one scholar boldly stated, not only social science but "*thinking* without comparison is unthinkable" (Swanson 1971, 145, emphasis added). We usually understand things in reference to other things. As social scientists we want to know if patterns exist across the social groups or individuals that we study, and pattern finding of course necessarily involves comparing. In fact, one of the founding fathers of sociology, Emile Durkheim, claimed that "comparative sociology is not a particular branch in sociology; it is sociology itself" (1938, 139).

Where better then to highlight the salience of comparison than in a cross-national setting? The all-too-familiar notion of "culture shock" is based on the premise that countries differ from one another. In our own familiar environments we tend to take things for granted. We often conflate the way things are around us with the way the world *should* be in general. But almost as soon as we get off the plane in a foreign country, we are nudged into questioning this assumption. As many students who return from foreign exchange trips reveal, an extended period of time in a foreign