Paratyping: A Contextualized Method of Inquiry for Understanding Perceptions of Mobile and Ubiquitous Computing Technologies

Gillian R. Hayes

University of California, Irvine, Irvine, CA, USA

Khai N. Truong

University of Toronto, Toronto, ON, Canada

RUNNING HEAD: A CONTEXTUALIZED METHOD OF INQUIRY FOR UNDERSTANDING PERCEPTIONS OF MOBILE AND UBIQUITOUS COMPUTING TECHNOLOGIES

Corresponding Author's Contact Information:

Gillian R. Hayes
University of California, Irvine
Department of Informatics
Irvine, CA 92697-3440
United States

Brief Authors' Biographies:

Gillian R. Hayes is an Assistant Professor in the Department of Informatics at UC Irvine. Her primary research interests are in human-computer interaction and ubiquitous computing, specifically in the areas of record-keeping and identity management systems and support for education and healthcare. **Khai N. Truong** is an Associate Professor in the Department of Computer Science at the University of Toronto. His research interests are in human-computer interaction and ubiquitous computing, specifically in the areas of assistive technologies, mobile interaction techniques, and sustainability.

ABSTRACT

In this paper, we describe the origins, use, and efficacy of a contextualized method for evaluating mobile and ubiquitous computing systems. This technique, which we called "paratyping," is based on experience prototyping and event-contingent experience sampling, and allows researchers to survey people in real-life situations without the need for costly and sometimes untenable deployment evaluations. We used this tool to probe the perceptions of the conversation partners of users of the Personal Audio Loop, a memory aid with the potential for substantial privacy implications. Based on that experience, we refined and adapted the approach to evaluate SenseCam, a wearable automatic picture-taking device, across multiple geographic locations. We describe the benefits, challenges, and methodological considerations that emerged during our use of the paratyping method across these two studies. We describe how this method blends some of the benefits of survey-based research with more contextualized methods, focusing on trustworthiness of the method in terms of generating scientific knowledge. In particular, this method is a good fit for studying certain classes of mobile and ubiquitous computing applications but can be applied to many types of applications.

CONTENTS

1. INTRODUCTION

In the past decade, we have witnessed a rapid proliferation of small, digital, ubiquitous recording technologies, including everything from camera-phones to sensor networks. Researchers have been building advanced functionality as well as the enabling technologies that underlie these applications. Examples of these technologies from the research literature often fall under the broad category of "capture and access" applications (Abowd and Mynatt 2000; Truong and Hayes 2009) and include the Personal Audio Loop (Abowd *et al*. 2005; Hayes *et al*. 2004), SenseCam (Hodges *et al*. 2006), Scribe4Me (Matthews *et al*. 2006), WWIT Recorder (Vemuri *et al*. 2004; Vemuri *et al*. 2006). At the same time, researchers have been examining how novel recording technologies can be used to support a variety of human needs (*e.g.,* (Barreau *et al*. 2007; Berry *et al*. 2009; Sellen *et al*. 2007; Yeh *et al*. 2006); for a more detailed review, see (Truong and Hayes 2009)). These projects tend to be grounded in understanding of the needs of potential users, often through fieldwork, interviews, and surveys. They then are often evaluated through laboratory studies of use or field trials. However, a critical examination of these projects, including many of our own, highlighted concerns that can arise in use and are not typically well addressed during early design efforts in the creation of ubicomp systems.

Responses to novel ubicomp technologies depend on many factors, including the social context of the specific interactions and the environment at the moment of their use. In addition, the mobile and often hidden nature of ubicomp systems requires that special care be taken to understand the responses to, concerns of, and preferences of those individuals whom a user of these technologies might encounter and therefore impact with the system. Users may have legitimate interests in using these tools, may not judge their use as particularly sensitive, and so on. However, these needs and considerations may be in direct opposition to the perceptions and preferences of those with whom they are interacting who might prefer the tools not to be used. This issue is particularly salient in the examination of capture and access tools, in which portions of the environment and those people in it may be recorded, but also applies to a variety of other ubicomp tools, including location and context-aware systems and so on.

These factors can often be explored through in-depth field deployments. However, such research requires working systems and often comes too late for the kind of early design intervention required to make a difference in the fundamental nature of these systems. In addition, conducting this kind of research through real-life use of ubicomp technologies can put people at risk through exposure to situations in which unpredictable reactions of others can damage the relationship between these groups or cause other disruptions. Likewise, these kinds of real-life deployments can result in the collection of data both the users and others around them who might actually object, thereby putting them at unnecessary risk for the goals of the research study, an important consideration in terms of both research ethics and the approval of research activities by regulatory boards. The very unpredictability of these settings makes them interesting for study and important to the design process.

Thus, in response to this need for a lightweight tool for conducting contextualized design research, we created an event-contingent experience sampling procedure focused on imagined uses of technologies in real situations we call paratyping. Event-contingent sampling procedures are those initiated by the occurrence of a specific event (Wheeler and Rois 1991). In this paper, we present our experience with developing and refining paratyping. We describe its usage in assessing responses to encounters with two types of recording systems: PAL and SenseCam. We demonstrate how the empirical data collected through this method enables adjustment of features early in design, prior to the hardening of particular technologies and without putting users and others around the technologies at undue risk. Finally, we discuss the overall trustworthiness of paratyping as a relatively contextualized but non-invasive method for gathering empirical data in terms of credibility, transferability, dependability, and confirmability.

2. BACKGROUND

There is substantial background work in HCI on the idea of *in situ* evaluations of systems within the kind of complex experiences imagined by ubiquitous computing. In particular, the Experience Prototypes used by Buchenau and Suri (2000) and Wizard-of-Oz techniques used in mobile settings (Li *et al.* 2004) are aimed at evaluating mobile technologies within iterative development processes. However, these approaches base their evaluation on reproduced or simulated experiences with the aim of testing the technology, instead of focusing on experiences with novel technologies in specific lived experiences.

Diary studies (Rieman 1993) and experience sampling (Larson and Csikszentmihalyi 1983) are popular HCI practices that do focus on gathering data close to the actual lived experiences of respondents. In both types of studies, participants may be asked periodically or randomly to document their experiences. In diary studies, often it is information about a particular experience that is gathered and preserved. For example, a person might be asked to document every time he or she switches tasks in an attempt to understand interruptions in the workplace (Czerwinsky *et al*. 2004). Diary studies have several possible limitations. First, the times at which participants choose to journal may not be situations of interest to the researchers. To counter this, participants are often asked to recall specific experiences, which may be of greater interest to the researchers. However, this method often introduces a memory bias, in which participants only document those experiences they remember rather than all those of interest to the researchers. In experience sampling, often participants are asked to provide reactions to a random simulated situation or to simply do so on a random or predetermined schedule. Consolvo *et al*. noted that the random simulated requests that a device in an experience sampling study made to participants were in various occasions implausible from a social standpoint (Consolvo *et al*. 2005). For example, the device would simulate a person asking for the participant's location when the participant felt that a person would have not done so in reality. The event-contingent nature of paratyping, by contrast, makes the responses potentially more salient or at the very least the requests for information more probable.

The use of simulated experiences in Experience Prototypes and Wizard-of-Oz, imagined technologies in Diary Studies and Experience Sampling, and abstract concepts in other survey methods all suffer from an inherent challenge when considering ubicomp

technologies, particularly in relation to designing around concerns for privacy, control, and power. Researchers have long recognized that people often take a deontological stance when artificially probed on opinions and preferences on privacy, both in reference to organizations (Berednt *et al*. 2005; Nguyen *et al*. 2008) and in interpersonal relations (Goffman 1959). Thus, the challenge becomes to understand the various, often conflicting, views as stated in the abstract, views as stated in the specific, and behaviors as practiced. This kind of understanding demands the use of different methodological approaches, particularly during the design process.

Stated preferences do not always match with everyday behaviors for a variety of reasons. However, people do tend to have a nuanced sense of privacy balance and highly developed practices to ensure appropriate distribution of power and control in interpersonal relations (Altman 1975). In social settings, individuals may choose certain paths of behavior to avoid conflict or in response to overriding social goals. Likewise, they may choose to give up control of their own data and negotiate, on the fly, more open responses to data typically considered more private in light of the social good and a communitarian ethic (Etzioni 1999).

Thus, in this work, we were particularly concerned with developing and testing a method that focuses on the perceptions of novel mobile technologies *during lived experiences*. It was our hope that by situating responses within deeply contextualized experiences, we might at least partially bridge the gap between action and thought with regard to mobile recording tools.

3. THE PARATYPING METHOD

We developed and evaluated a contextualized method for inquiry about novel technologies. We called this method "paratyping" after the concept of a paratype: a simulation or model of interaction (*-type*) with a technology evaluated alongside (*para*) real world experience" (Iachello *et al.* 2006). This method focuses on gathering feedback from two types of participants, who we call *proxies* and *survey recipients.* The survey instrument used is distributed by individuals we call "proxies" because they serve as proxy users of the technology being investigated. These individuals act as if they are using the device being interrogated—but do not actually use it—and distribute surveys as they go about their daily activities to anyone fitting the criteria of the research project at hand. In all of our studies, proxies were also instructed not to distribute the survey to the same individual more than once to reduce sampling bias, although this solution is optional, and statistical measures during analysis could accomplish some of the same control.

Using this method, the survey itself can be adapted to fit the needs of the specific research project. However, it should instruct survey recipients to suppose that the proxy has been using the device in question, and probe their opinions and feelings about that device in *that specific situation*. In this respect, this technique is similar to Critical Incident techniques developed in the context of workplace psychology (Flanagan 1954). This procedure allows the researchers to situate participant response in the experience the person just had, with a specific user, in a specific location. This contextualization should reduce recall errors and hypothetical answers. Although the survey was administered by human proxies in both of the case studies we present in this article, it is not part of our definition of paratype. The surveys could be administered in some automated manner

given the right research questions and technologies. The term paratype only refers to introducing simulated interaction with a certain technological artifact within a specific setting of real social action, and documenting the effects of this combination (Iachello *et al.* 2006).

Regardless of who or what plays the role of the proxy, the proxy's role in this method is much more than simply administering a survey. Any research assistant, auto-emailer, or website could serve that function. What is of methodological interest here to the HCI community, however, is the primary function of the proxy: to *create the technological instance* on which the researchers want feedback. In our case studies, the proxies accomplished this act of creation by behaving as if they were users of the technology in question. Thus, by going about her daily activities, the proxy systematically creates the experiences to be examined with the help of the description of the application and, if requested, a demonstration of the working device.

<<Figure 1 About Here>>

3.1 Designing a Paratyping Study

A paratyping survey should be composed of two parts, linked by a unique number (see Figure 1). The first part (on the left) is completed by the proxy, with information about the "salient elements of the social setting" (Goffman 1966). The portion on the right then is detached and given to the survey recipient following a qualifying interaction, as defined by the inclusion criteria for the particular study. The surveys should be designed to be self-explanatory, and contain descriptions of the system or device being queried, the research ethics, and relevant questions. The descriptions of the technologies should be validated prior to the study to ensure they are pertinent and sufficient. Typically, the systems themselves are not operated using this method. It is most appropriate, after all, for those types of tools for which long-term operation is technically or ethically infeasible given the research questions. Participants should then be asked to complete the survey immediately if possible, to increase recall accuracy. Depending upon the logistics of the study, the requirements of the local ethics boards, and other pragmatic issues, the survey portion of the card can be return-addressed on the backside and mailed back at the participant's convenience or collected by the proxies at the time. The questionnaire should be short (less than 10 questions, all of which can be answered by circling, checking, or writing just a few words) to encourage rapid completion and temporally closely related responses.

This model of research requires training for each proxy. In practice, in our experience, these training sessions can be completed in approximately an hour. During these sessions, the researchers should explain the purpose of the study and then describe the technology and demonstrate its uses. Of course, the procedure of distributing surveys should be explained, including the inclusion criteria for survey recipients (*e.g.,* anyone over 18 years old who has not received a survey before and with whom the interaction lasted at least 30 seconds). Finally, each proxy should be trained and tested to ensure they understand the technology and the study procedures.

The demographics of the proxies are likely to influence the demographics of the respondents, in terms of age, socio-economic class, education, *etc.* We call this issue social network bias and note that it can strongly influence the results. Additionally, although the protocol calls for distributing the survey after every qualifying interaction, sometimes a proxy might choose not to distribute surveys in some situations. To mitigate these issues, researchers should recruit as large and diverse a group of proxies as their resources allow and target population demands. Proxies can be recruited using a variety of methods. We have successfully used snowball sampling from our close social networks, posting to mailing lists, and posting to short-term job opportunity websites (*e.g.,* Craigslist), but recruitment methods for these individuals should be tailored to the specific research questions being addressed in any given project.

Interviews are not an essential element of the paratyping method, but they can be used in conjunction with the surveying practices to add substantial depth to the findings. Interviewing both proxies and survey respondents provides a more detailed and nuanced understanding about why participants responded the way they did in the moment than the surveys alone. Additionally, respondents are asked to complete the survey as close to in the moment as possible. Over time, they may reflect on their answers and wish they had responded—or if the actual technology had been in use, acted—differently. Interviewing

respondents later in the process gives them a chance to explain their responses as well as any changes they might have made given further reflection. This act of reflection can be particularly informative in terms of how the technologies might be designed, preventing the kind of initial enjoyment and eventual distaste or even disgust that so many novel technologies engender in use. Thus, researchers may wish to include three types of interviews:

- *Ongoing proxy interviews* enable researchers to collect completed surveys, follow up with training and questions, and provide ongoing compensation for participation in the study. These meetings provide the opportunity to engage the proxies in questions focused on their own comfort with the technology, their experiences with their interaction partners about the technology and about the study, and so on.
- *Proxy exit interviews* allow proxies to reflect on the technology in question as though they had been using it in a variety of settings as well as to describe any experiences with the research procedures that may bias the responses collected from those surveys they distributed.
- *Survey respondent follow-up interviews* further probe their responses to the survey and their attitudes towards the specific technology in question and other related issues, systems, and designs.

Researchers should take detailed field notes during all of the interviews and record and transcribe them when possible. These data can then be coded alongside relevant theoretical frameworks, in connection with the quantitative results of the surveys, or in an open process that allows for the emergence of new themes.

<<Figure 2 About Here>>

3.2 Paratyping in Use

Our early experiences with paratyping were all in the United States, in one language. As both we and other researchers became interested in using this method in other countries, we questioned whether the paratyping method would transfer well into other national and cultural contexts. Of course, procedures must be conducted in a language comprehensible to the participants. Validation of survey instruments in multiple languages is a complex and difficult process, one that can take years in other disciplines (Li *et al.* 2001). In terms of paratyping, because the method can be used with previously validated instruments or with one of the researchers' design, we leave the decision about validation of multiple languages up to the discretion of those researchers. In our own experiences, we used short surveys (less than 10 questions) of our design, and when necessary translated those into other languages using a translator or co-researcher who is a native speaker. Additionally, surveys were checked for any local colloquial language concerns. In multi-language contexts, as in one of our studies, the native language of the proxies should also be taken into consideration, but surveys can be provided in multiple languages to multi-lingual proxies with appropriate training. Minor differences in language introduced through translation can result in different understandings of and responses to survey questions and must be monitored closely. In addition, subtle differences in the cultures in which these studies are conducted can result in changes in

the ways in which the surveys are executed. For example, in some cultures and nations, the distribution of such an instrument may not be appropriate at all or must be handled in a way that demands special training of the proxies, return of the surveys directly to the proxies rather than mailing them in, and so on. These changes—particularly when conducting a study in multiple places simultaneously—can introduce new biases, which must be considered both at study design and analysis time.

Paratyping emerged as a method, in part, in response to the ethical concerns of the undue risk and burden to users and their friends, family, and acquaintances in using prototypes of ubicomp technologies. However, new ethical concerns can arise with such a method, in particular around the issue of consenting to be part of the research. During the initial survey portion of the study, respondents only interact with the proxies not with the researchers directly thus begging the important question of who informs potential research participants about the study and who collects (and documents) consent. Furthermore, although we recommend that the instrument itself always include contact information for the researchers in the case of any questions, the additional time and effort required of respondents to do so may override any desire they have to learn more about the study. Although they can always choose not to mail in the response card, this hard line between participating or not is rather extreme compared to the ease of being able to ask a researcher a quick question about the study and then proceed with participation as would occur in an interview study. Proxies may be instructed not to answer such questions to increase the standardization across the study, but questions of whether this kind of instruction can be enforced and whether it should be delivered in the first place are complex and not yet fully explored.

Ethical behavior is important in all forms of research, and thus, most institutions have substantial rules, regulations, and procedures in place to protect the individuals who participate in research studies. Although institutional approval and ethics are not one in the same, they are intrinsically tied in practice. In all of our studies, the researchers were granted waivers of documentation of informed consent (*e.g.,* signing a consent form) for the survey portion of the research. Because the surveys are anonymous, except when respondents provide additional contact information for follow-up interviews, the documentation of consent would actually have put the participants at greater risk by identifying them and storing their signatures on file. The majority of the institutions involved in approval of these studies also considered proxies to be research subjects. One institution, however, required that they take human subjects protections training and be treated as research staff, potentially biasing their responses to interviews and surveys themselves. The informed consent procedures for those participants willing to be interviewed after responding to a survey were similar to those for interview protocols generally at each institution. As in the PAL study, the researchers at both institutions worked closely with the governing boards to ensure an appropriate solution was reached.

4. CASE STUDIES

A method can often be understood best through its application. To that end, in this section, we describe two case studies in which we used paratyping: studies of the Personal Audio Loop (PAL) and SenseCam. Paratyping is particularly useful in the evaluation and design of applications for which a prototype system cannot be deployed due to technological infeasibility, social and ethical concerns, or both. These case studies

are representative of these types of applications in that both raise substantial social and ethical concerns—both are mobile recording applications that are nearly invisible to anyone but the user. Only PAL was technologically challenging for long-term deployment, however. At the time of the evaluation described here, SenseCam had already been used extensively in other deployments. Thus, these case studies demonstrate how this method can be used both for design, as in its initial purpose, and evaluation. We refer to these case studies extensively in our discussion, thus warranting a brief overview here. The results of those two studies have been previously reported separately. Thus, here we do not focus on those results in particular, but rather we use these cases to discuss the successes and challenges with using the paratyping technique in the evaluation of mobile and ubiquitous computing technologies.

4.1. The Personal Audio Loop (PAL)

The Personal Audio Loop (PAL) is a short-term mobile audio memory aid that runs on a mobile phone (Abowd *et al*. 2005; Hayes *et al*. 2004; Patel *et al*. 2008). Users can replay sounds heard in the recent past up to a defined maximum time. Audio older than this retention time is automatically deleted one second at a time. PAL records at all times that the user is not using the phone for telecommunication. Users rewind, replay, and pause recordings through a few buttons on the side of the phone, making the application operable one handed. Although participants in various studies of PAL recognized its usefulness nearly universally, they also raised concerns about how it functioned. PAL records continuously, unattended, and unnoticed by the user and potential conversation partners or bystanders. The primary concerns related to the impact on the privacy and control of data for conversation partners and unrelated third parties (*e.g.,* passersby). Further concerns related to the social appropriateness of using the application, regarding both the immediate disruption of interpersonal interaction and long-term effect on social relationships.

<<Figure 3 About Here>>

To provide a better understanding of the situations in which its usage is acceptable and the parameters of recording which influence its acceptability, we designed and executed the paratyping technique to sample people's immediate reaction towards PAL during their potential first encounters with such a device (Iachello *et al.* 2006). The paratyping study of PAL included three proxies: two females and one male between the ages of 27 and 31, all HCI graduate students and researchers working at the same institution, living alone or with a partner. Of 45 distributed surveys, 41 resulted in usable responses (17 from females). Most surveys were completed immediately, and 9 were mailed back afterwards. Of the 41 respondents, 24 were in IT or research occupations (students, research scientists, university professors, *etc.*). The remaining respondents ranged across professions, including: teachers, designers, hairdressers, managers, attorneys and business owners. Respondents spanned all age groups between 18 and 60 and over. However, age distribution was biased towards the younger age groups (the median age group was 30–39), reflecting the age group of the proxies.

The detailed results of this study can be found elsewhere (Iachello *et al.* 2006), however it is worth describing the overall tenor of these results briefly. Participants noted

that awareness and greater understanding of PAL was important to allow "boundary-setting" to occur. By boundary-setting, participants in these studies were describing the process by which they negotiate with those around them the appropriateness of recording audio and using it for particular purposes in a variety of contexts. Informal social boundaries, such as avoidance of certain topics of conversation with certain friends or keeping a relatively large physical distance from other friends, are one way that people can negotiate complex settings and avoid offense in the absence of formal policies. Considerations of boundaries also brings to mind the dynamic and dialectic processes of negotiating the boundary between public and private as described initially by Altman (Altman 1975) and later taken up by Palen and Dourish in regard to information privacy in particular (Palen and Dourish 2003). Survey respondents were not concerned as much by retention time as with potential misuse of the recordings. They also stated that they would have rarely asked to delete a recording after the fact. We did not ask if they would like the recording deleted, merely whether they would ask for it to be deleted. Thus, this result tells us more about their feelings of agency and the level of concern they might have than about general preference for deletion. Overall, the results of this study suggest that traditional privacy guidelines and policies may not be appropriate nor sufficient for the development of these types of personal recording tools and that designers should focus on the purpose of use of information and interpersonal dynamics instead. Furthermore, these results suggest that even when concerns may be present, those being recorded do not necessarily believe they can or should—within the restrictions of societal norms, personal relationships, and even legal guidance—prevent or alter a user's recording with these tools.

In the PAL project, we had already conducted a laboratory study, diary study, and numerous interviews of potential users to understand the challenges and opportunities of near term mobile audio recording. Although these studies were informative, and we by no means advocate the removal or replacement of these methods in the suite of ubicomp design and research methods, the addition of paratyping enabled us to uncover new findings that were not visible through these other methods. For example, in interviews, we probed potential users about instances of use of PAL. These kinds of interviews either require participants to summon their own examples of use, which is a creative and challenging exercise in an interview setting or require them to respond to potential scenarios generated by the researchers. In the latter case, participants had to reflect on situations that were not necessarily regularly encountered, and it would be difficult to create a set of scenarios representative of all possible conversational situations. With paratyping, we were able to capture a variety of situations, many of which were outside our initial conceptions of when and how PAL might be used, such as a salesperson recording a conversation with a customer and replaying it for a trainee. In an interview study—or even a long-term deployment—we might have gathered information from the salesman or even the trainee, but the views of the customer who was recorded would almost certainly not have been included in the dataset.

4.2. SenseCam

As a second example, SenseCam[1] is a wearable automatic picture-taking device (Hodges *et al*. 2006) that has been used in a wide variety of situations to support diverse human needs—*e.g.,* (Doherty and Smeaton 2009; Fleck and Fitzpatrick 2009; Lee and Dey 2008; Yang *et al*. 2006). SenseCam is approximately the size of a deck of playing cards with battery life and storage capacity of a day. The device is designed to be worn and takes pictures every 30 seconds or whenever it senses a likely situation of interest. The sensors used are relatively simple and include such things as accelerometers for detecting movement of the user, light sensors for detecting movement around the user, and a microphone for detecting but not recording audio. A simple button allows pausing of the recording of images but no other interaction happens on the device itself. SenseCam images are reviewed on a desktop computer using software built in to the device. Researchers have conducted numerous studies of SenseCam use with patients with memory impairment—*e.g.*, (Hodges *et al*. 2006, Sellen *et al.* 2007)—in educational settings (Barreau *et al*. 2007), in business negotiations with blind users (Tjoa *et al*. 2006), and more. During previous studies of SenseCam, however, researchers reported that most of the people with whom the wearer interacted either did not notice the device or noticed it but comprehended neither its capabilities nor uses and that SenseCam users tended to view the device as a personal tool not requiring the permission of interaction partners to use it. Thus, in cooperation with SenseCam developers, we adapted the paratyping technique to assess the situations in which its usage is acceptable to these interaction partners and how its parameters for recording influence its acceptability (Nguyen *et al.* 2009). We conducted a larger scale study of SenseCam than PAL, spread across four geographic locations, to examine possible cultural differences and test the internationalization of the paratyping method.

The paratyping study of SenseCam included 19 proxies distributing nearly 700 surveys in total. This research was conducted in Toronto, Ontario, Canada (CAN, 3 proxies), Orange County, California, USA (US, 5 proxies), Cambridge, England (UK, 6 proxies), and Zurich, Switzerland (CH, 5 proxies). The proxies included 7 men and 12 women, aged 18 to early 60s with a variety of professions (*e.g.,* security guard, architect, caterer, student). We received 413 responses from 686 eligible encounters. The survey was not returned in 233 instances and not distributed in 40 encounters making a 64% response rate. Respondents ranged across professions, including: teachers, actors, designers, attorneys, realtors, receptionists, engineers, managers, restaurant staff, entertainers, and more. Of 413 respondents (205 female, 205 male, 3 undisclosed), 206 were under 30; 87 in their 30s, 60 in their 40s, 31 in their 50s, and 25 were 60 or older. Four participants did not provide age.

As in the PAL study, the detailed results of this work can be found elsewhere (Nguyen *et al.* 2009). However, for the purposes of discussion, it is again worth describing a quick summary of these results. Our results indicate that individuals across all four countries involved in this study engaged in complex – though often rapid – reasoning and decision-making about SenseCam. People assessed, understood, and responded to SenseCam by drawing from their own personal values and beliefs, institutional and societal norms and customs, and their understanding of the technological

[1] SenseCam has recently been released as a commercial product, known as ViconRevue. However, when this work was conducted, SenseCam was still available only as a research tool, and thus we reference SenseCam not ViconRevue throughout this work.

features of SenseCam. The results of this study indicate that people would tolerate potential incursions from SenseCam for particular purposes. Furthermore, they would typically prefer to be informed about and to consent to recording as well as to grant permission before any data are shared. These preferences, however, are unlikely to instigate a request for deletion or other action on their part. There were some cultural differences observed among the responses, in particular in relation to comparisons of SenseCam to CCTV (UK) and concerns about attractiveness and aesthetics (US).

<<Figure 4 About Here>>

Like in the PAL study, using the paratyping method to evaluate SenseCam was most useful in terms of what it told us about the interaction partners' views and preferences of the device. In this study, half the proxies carried or wore the device while the other half relied only on the written description as in the PAL study. No statistically significant differences were found in the responses to surveys distributed by SenseCam carriers as opposed to those with only the written description, and so we do not advocate one approach or the other. However, in interviews with respondents who had been shown the device, we were able to get some insight as to their initial reactions to its form factor. No one recognized it to be a camera with explanations for its actual purpose as far ranging as a personal air purifier. The reactions to the device once it was explained were diverse, and both the initial impressions and responses from these individuals as their knowledge grew would have been difficult, if not impossible, to get in other ways.

5. DISCUSSION

Over the course of five years and multiple studies using this method, we have continued to refine it based on those elements we believe to be working well and those that limit its overall trustworthiness as a method for near-naturalistic inquiry (Lincoln and Guba 1985). In this section, we describe how the method can be judged in terms of trustworthiness. Trustworthiness is a useful consideration for empirical methods that are largely devoted to concerns around design, as opposed to, for example, generalizable models for human behavior. The paratyping method is most useful as a design method to be interjected within iterative cycles of design and development of ubicomp and other hard to prototype systems. In these cases, the ultimate goal is the design of workable, usable, and useful systems, making trustworthiness of the process and design considerations to emerge from it most important for generating knowledge in these settings. For the purposes of this discussion, we use Stringer's notion that scientific trustworthiness stems from four related issues: credibility, transferability, dependability, and confirmability (Stringer 2007, p.57). Credibility here refers to the integrity of the study, transferability to the ability to transfer the results to another context—notably not generalizability, dependability to the clarity in description of the features of the study, and confirmability to the evidence that the procedures described actually took place. In this section, we describe some of the features, caveats, and limitations we uncovered through our use of this method in relation to these four issues. Additionally, we offer a discussion of the research ethics of using the paratyping method.

5.1. Credibility

The paratyping method allows for collection of empirical data that would be impossible to gather in a typical survey or in a laboratory setup. It could be possible to gather situated responses using shadowing or other in-depth observational techniques, and we generally recommend their use when the research questions demand this level of inquiry. However, paratyping enables contextualized data collection in an arguably less intrusive and definitively less expensive manner for studies for which that kind of in-depth, long-term engagement is either not feasible or unnecessary.

Paratypes combine event-contingent experience sampling with experience prototyping. Thus, they are particularly useful for evaluating interactions and experiences in which reference to concrete instances of everyday life are needed. Often, the kinds of contexts that might arise to make use of a particular technology useful, enjoyable, inconvenient, or problematic are not easily predicted (and thus hard to survey) and are even harder to simulate (and thus cannot easily be tested in the laboratory). Furthermore, absent the details of a particular experience, it may be hard to articulate just what is problematic or helpful about imagined uses of technology. For example, in the PAL study, we recognized that people often do not know how to articulate how often or when they are in "confidential" conversations nor even what makes them confidential except by example. Although justifications and perceptions of use given at the time of the experiences are subject to rationalization and performance for the researchers on the part of the survey respondents, these issues are somewhat different than those that arise when a respondent is attempting to describe a situation they may not fully recall and from which they have had sufficient distance to contemplate and rationalize their responses over time. With paratyping, the goal is to capture those "knee-jerk reactions" and situate descriptions of respondent behavior and preferences in the moment without intensive ethnographic or shadowing methods.

One threat to the credibility of this method is the introduction of bias through sampling. Necessarily, the life experiences of the proxies influence the demographics of those sampled in the survey portion of the study. Additionally, sample selection bias is at issue with the recruitment of follow-up interview participants through the collection of voluntary contact information on the surveys themselves. We first saw the sampling bias of survey respondents in the PAL study. To control for this issue, in the SenseCam study, we attempted to recruit a broader range of proxies. However, the demographics of survey respondents still correlated to those of the proxies. Thus, researchers using paratyping must be cognizant of the likely sampling effects they introduce with the recruitment of proxies. This issue can be a benefit if the study is meant to target a very specific population (Milardo 1992).

Additionally, proxies do not always distribute surveys when an encounter qualifies for it, an issue that first emerged in the PAL study but has been seen in other studies since. For example, one proxy described discomfort with distributing surveys at a medical visit or when in a hurry. These settings may be exactly the ones about which the researchers want feedback, but they were not represented in the data. Thus, in the SenseCam study, we explicitly asked proxies to detail such encounters using the left side of the survey instrument without distributing the right side to the interaction partner. Proxies were also compensated for returning these surveys even if they did not distribute the respondent half. Although this solution means that the interaction partner's perceptions are still unrepresented in the data, we were able to account for those settings

with at least partial data. These situations can then be probed in more detail in interviews with proxies and respondents, a situation that is workable if not ideal.

Finally, just as in survey work in which respondents hope to please the researchers, respondents in these cases may be attempting to please the proxies. Although researchers using paratyping can control for this concern to a degree by asking respondents to mail the response card, they may not entirely trust that the proxies will never see the information. Thus, noting that they are concerned about the recording or do not trust a proxy to record and save media about them may be problematic for the respondents socially. In follow-up interviews with survey respondents, researchers can—as we have done in the past—attempt to unpack these complicated considerations and offer to the respondents to change their answers if they like. However, this solution would capture only a minority of survey respondents and may not address all of the bias encountered. Thus, as in all self-report survey research, researchers must take the self-report to be only a portion of the story and examine the issues at hand with a variety of approaches.

5.2. Transferability

To date, our evaluations using paratypes have only focused on mobile and ubiquitous computing technologies. Specifically, we have focused on systems that collect information at unexpected times or even continually. However, we believe these techniques can also be helpful for other mobile and ubicomp solutions, such as those that provide information when needed (*e.g.*, portable guides) or those in which interaction is embedded in un-planned social practice or everyday routine, such as home communication systems (Nagel *et al.* 2004). For example, Ollila *et al.* discuss the use of paratyping to gather feedback from players of pervasive games, which put them in new kinds of social situations or require them to change their behavior (2008). These kinds of applications and systems often have high prototyping costs. Furthermore, their portable nature makes it simultaneously important and difficult to probe their design and use in context.

We do not argue for the results of paratyping as always generalizable nor predictive. These results do, however, provide researchers with highly situated data that can help designers to reason about what it was about the context that may have contributed to the character of the responses. This kind of reasoning supports cross-contextual transfer of results and design implications. Furthermore, the method itself is highly extensible and transferable. Thus, as more researchers make use of the method, further comparative analysis can be undertaken.

5.3. Dependability

Dependability centers on the notion that all procedures required of systematic research have been followed and clearly articulated. As a primarily survey-based method, the dependability of paratyping is arguably quite high. All survey respondents theoretically receive the exact same instructions, because they are provided in written format with a note that follow-up questions should be directed to the researchers not the proxies. As mentioned in the paratyping method section, in practice, no one has ever requested further information in any paratyping study we have conducted. The written descriptions of the devices should be tested repeatedly before deploying the instrument to

insure comprehensibility but questions may remain, and the overhead of contacting a researcher, as discussed, may simply be too high, leaving respondents to query the proxies regardless of instructions.

The use of the proxies inherently threatens dependability of the method. First, proxies must be trained to conduct the research properly. Every individual learns differently, and the training may even be delivered slightly differently for different proxies. Second, proxies, in their roles as research participants, tend to be compensated for their participation. Any compensation scheme will influence the way the proxies behave. It is possible—perhaps even likely—that proxies in a compensation scheme that privileges the distribution of surveys will give them out to people with whom there was no qualifying interaction. On the other hand, a compensation scheme that rewards the return of surveys might encourage proxies to fake data and return the surveys themselves or to coerce their interaction partners into returning surveys.

As mentioned when describing the method initially, the proxies do not have to be humans, although they have been in our studies. It is possible to use automated means for distribution of these surveys depending on the technology and issues being probed. However, use of non-human proxies would likely introduce other concerns. It is important for researchers to develop good relationships with human proxies when they are used and ensure their training includes an understanding of the need for dependable research results. Furthermore, regardless of the type of proxy being used, the process by which they proxies are instructed to distribute surveys must be closely analyzed and standardized when possible to ensure that these issues are reflected in the way the resultant data are understood.

5.4. Confirmability

Researchers, reviewers, and ultimately readers examining the results of a particular study should be able to confirm the veracity of a study. Audit trails and collected data provide this evidence in many cases. In paratyping, the majority of the data are survey responses, which are easily documented. Ideally, there are also transcripts of audio-recorded interviews with proxies and survey respondents as well as documented training scripts. This level of detailed documentation inherent to paratyping makes its confirmability relatively high compared to many other methods.

This kind of evidence, however, can be incredibly challenging to present in the form of a conference or journal article and is often only to be found in the hundreds of pages of a doctoral dissertation. The ACM's new practice of allowing the addition of datasets to the online archive as well as the policies from the US National Science Foundation around the publication of data may support the practice of publishing these kinds of data[2]. However, with paratyping as with any method, often the collection of the data can be difficult enough that the researchers do not want to publish their data for fear that other researchers might publish re-analyses of the results before those who initially collected the data have had a chance to fully complete their own efforts. Thus, policies must be developed around these practices to ensure that confirmability can be met for paratyping

[2] http://www.nsf.gov/news/news_summ.jsp?cntn_id=116928, http://tochi.acm.org/authors.shtml

and other methods of human inquiry without diminishing the data collection, analysis, and publication practices already in place.

6. CONCLUSION

People's perceptions of and reactions to novel recording technologies depend on many factors. The context of the specific interactions and experiences at the moment of recording can substantially influence reported perceptions about the technologies in the moment, indicating both the necessity for designers to include flexible dynamic management of recording features and the need for design methods that explicitly interrogate responses to early prototype designs in a variety of situations and settings. Understanding perceptions of and reactions to novel mobile and ubiquitous computing technologies requires contextualization of the experience in which the technologies are encountered. However, it is difficult, and often disruptive or otherwise problematic, to address these questions through extensive long-term deployment exercises, shadowing, and other fieldwork oriented methods.

Over the last five years, we have developed and refined a contextualized method for evaluating mobile and ubiquitous computing systems, called paratyping. This method relies on the notion of proxy encounters with the novel technology that situate survey responses in the use of that technology in a particular time and place. In this paper, we demonstrated the use of paratyping in two case studies, which we have used to assess people's reaction to encounters with the Personal Audio Loop and the SenseCam. The results from these two case studies demonstrate how paratyping can be used successfully to query the potential for adoption and use of mobile and ubiquitous recording technologies in a contextualized manner without the need for extensive deployments. Data gathered through this method demonstrate how the technological capabilities contribute to the acceptability of audio and image recording, how different cultures respond to image recording, and generally how people feel in specific situations about the potential for recording through these novel technological systems. Our results further confirm the highly contextualized nature of responses to these types of technologies, necessitating use of methods like paratyping.

The method and two case studies presented in this work lay the groundwork for future studies using this method as well as further methodological development of techniques and tools for situated inquiry in HCI. When deigned properly, paratyping studies can offer trustworthy results in terms of credibility, transferability, dependability, confirmability. This method can be usefully applied to other projects, particularly those related to mobile and ubiquitous computing. Of course, those threats to trustworthiness described in the previous section must be considered when designing a paratyping study. Additionally, we leave open for future work the adaptation of this method to other research projects.

NOTES

Acknowledgments. We thank Giovanni Iachello, Molly Stevens and Gregory Abowd for their help in designing and executing the original paratyping study with the Personal Audio Loop. We also thank Shwetak Patel for writing the PAL application, Jens Großklags, Elaine Huang, Kris Nagel, Delphine Nain and Jay Summet and the CHI reviewers for their thoughtful comments on a paper focused on the PAL study. We thank David Nguyen, Gabriela Marcu, James Scott, Marc Langheinrich, and Christof Rudoner for their help in designing and executing the paratyping study with SenseCam. We also thank Paul Dourish, Irina Shklovski, and the members of the STAR Group for their comments on that work. Additionally, we thank all study participants for their time and feedback.

Support. The paratyping study with the Personal Audio Loop was supported in part by NSF ITR grant 0121661, the NSF GRFP, the MacArthur Foundation, Intel Corporation and Motorola Corporation and was covered under IRB protocol H04153 at the Georgia Institute of Technology. The paratyping study with SenseCam was supported by NSF CAREER Grant 0846063 and support from Microsoft Research and was covered under IRB protocol 2007-6070 at the University of California, Irvine and 21027 at the University of Toronto.

Authors' Present Addresses. G.R. Hayes, Department of Informatics, Donald Bren School of Information and Computer Sciences, University of California, Irvine, Irvine, California, USA. E-mail:gillianrh@ics.uci.edu; K.N. Truong, Department of Computer Science, University of Toronto, Toronto, Ontario, Canada. E-mail: khai@cs.toronto.edu

HCI Editorial Record.

REFERENCES

Abowd, G.D., Hayes, G.R., Iachello, G., Kientz, J.A., Patel, S.N., Stevens, M., Truong, K.N. (2005). Prototypes and Paratypes: Designing Mobile and Ubiquitous Computing Applications. *IEEE Pervasive Computing, 4(4),* 67-73.

Abowd, G. D. and Mynatt, E. D. (2000). Charting past, present, and future research in ubiquitous computing. *ACM Transactions on Computer-Human Interaction, 7(1),* 29-58.

Altman, I. (1975). *The Environment and Social Behavior— Privacy, Personal Space, Territory, Crowding*. Monterey, CA: Brooks/Cole Publishing Company.

Barreau, D., Crystal, A., Greenberg, J., Sharma, A., Conway, M., Oberlin, J., Shoffner, M. and Seiberling, S. (2007).Augmenting Memory for Student Learning: Designing a Context-Aware Capture System for Biology Education. In *Proc. of the American Society for Information Science and Technology, 43(1),* 251–251.

Berendt, B., Günther, O., Spiekermann, S. (2005). Privacy in e- commerce: stated preferences vs. actual behavior. *Communications of the ACM, 48(4),* 101–106.

Berry, E., Hampshire, A, Rowe, J., Hodges, S., Kapur, N., Watson, P., Browne, G., Smyth, G., Wood, K., Owen, A. (2009). The neural basis of effective memory therapy in a patient with limbic encephalitis. *Journal of Neurology, Neurosurgery, and Psychiatry with Practical Neurology.* 164251.

Buchenau, M. and Suri, J.F. (2000). Experience Prototyping. In *Proc. of DIS 2000*, 424–433.

Consolvo, S., Smith, I., Matthews, T., LaMarca, A., Tabert, J., Powledge, P. (2005). Location Disclosure to Social Relations: Why, When, & What People Want to Share. In *Proc. of CHI 2005*, 82–90.

Czerwinski, M., Horvitz, E., Wilhite, S. (2004). A diary study of task switching and interruptions. In *Proc. of CHI 2004*, 175-182.

Doherty, A.R. and Smeaton, A.F. (2009). Utilising Wearable Sensor Technology to Provide Effective Memory Cues. *European Research Consortium for Informatics and Mathematics. ERCIM NEWS 76.*

Etzioni, A. (1999). *The Limits of Privacy.* New York, NY: Basic Books.

Flanagan, J.C. (1954). The Critical Incident Technique. *Psychological Bulletin, 51(4),* 327–359.

Fleck, R. and Fitzpatrick, G. (2009). Teachers' and Tutors' Social Reflection around SenseCam Images. *International Journal of Human Computer Studies, 67,* 1027-1036.

Goffman, E. (1966). *Behavior in Public Places.* Free Press.

Goffman, E. (1959). *The Presentation of Self in Everyday Life.* Anchor Books.

Hayes, G.R., Patel, S.N., Truong, K.N., Iachello, G., Kientz, J.A., Farmer, R. Abowd, G.D. (2004). The Personal Audio Loop: Designing a Ubiquitous Audio-Based Memory Aid. In *Proc. of Mobile HCI 2004*, 168–179.

Hodges, S., Williams, L., Berry, E., Izadi, S., Srinivasan, J., Butler, A., Smyth, G., Kapur, N., Wood, K. (2006). SenseCam: A Retrospective Memory Aid. In *Proc. of UbiComp 2006*, 177-193.

Iachello, G., Truong, K. N., Abowd, G. D., Hayes, G. R., Stevens, M. (2006). Prototyping and sampling experience to evaluate ubiquitous computing privacy in the real world. In *Proc. of CHI 2006,* 1009-1018.

Larson, R. and Csikszentmihalyi, M. (1983).The experience sampling method. *New Directions for Methodology of Social and Behavioral Science, 15,* 41–56.

Lee, M.L. and Dey, A.K. (2008). Wearable experience capture for episodic memory support. In *Proc. of IEEE Symposium on Wearable Computers,* 107-108.

Li, R.M., McCardle, P., Clark, R.L., Kinsella, K., Berch D. (2001) *Diverse Voices. The inclusion of language-minority populations in national studies: challenges and opportunities.* Bethesda, MD: National Institute on Aging and National Institute of Child Health and Human Development

Li, Y., Hong, J., Landay, J. (2004). Topiary: A Tool for Prototyping Location-Enhanced Applications. *Proc. of UIST 2004*, 217–226.

Lincoln Y.S. and Guba E.G. (1985). *Naturalistic Inquiry.* Beverly Hills, CA: Sage Publications.

Matthews, T., Carter, S., Pai, C., Fong, J., Mankoff, J. (2006). Scribe4Me: Evaluating a Mobile Sound Transcription Tool for the Deaf. In *Proc. of Ubicomp 2006,* 159-176.

Milardo, R. M. (1992). Comparative methods for delineating social networks. *Journal of Social and Personal Relationships, 9,* 447–461.

Nagel, K., Hudson, J., Abowd, G.D. (2004). Predictors of avail- ability in home life context-mediated communication. In *Proc. of CSCW 2004*, 497–506.

Nguyen, D.H., Hayes, G.R., Kobsa, A. (2008). An Empirical Investigation of Concerns of Everyday Tracking and Recording Technologies. In *Proc. of Ubicomp 2008*, 182-191.

Nguyen, D. H., Marcu, G., Hayes, G. R., Truong, K. N., Scott, J., Langheinrich, M., Roduner, C. (2009). Encountering SenseCam: personal recording technologies in everyday life. In *Proc. of Ubicomp 2009,* 165-174.

Ollila, E.M.I., Suomela, R., Holopainen, J. (2008). Using prototypes in early pervasive game development. *Computers in Entertainment, 6(2),* Article No. 17.

Palen, L. and Dourish, P. (2003). Unpacking "privacy" for a networked world. In *Proc. of CHI 2003,* 129-136.

Patel, S.N., Truong, K.N. Hayes, G.R., Iachello, G., Kientz, J.A., Abowd, G.D. (2008). The Personal Audio Loop: A Ubiquitous Audio-Based Memory Aid. *Handbook of Research on User Interface Design and Evaluation for Mobile Technology*, Joanna Lumsden, Editor, 1031 – 1048.

Rieman, J. (1993). The diary study: a workplace-oriented research tool to guide laboratory efforts. In *Proc. of INTERACT 1993 and CHI 1993, 321-326.*

Sellen, A., Fogg, A., Aitken, M., Hodges, S., Rother, C., Wood, K.R. (2007). Do life-logging technologies support memory for the past? An experimental study using sensecam. In *Proc. of CHI 2007*, 81-90.

Stringer, E.T. (2007). *Action Research (3rd Ed.).* Thousand Oaks, CA: Sage Publications.

Tjoa, A., Andjomshoaa, A., Karim, S. (2006). Exploiting SenseCam for Helping the Blind in Business Negotiations. *Computers Helping People with Special Needs,* 1147-1154.

Truong, K.N. and Hayes, G.R. (2009). Ubiquitous Computing for Capture and Access. *Foundations and Trends® in Human–Computer Interaction*, *2(2),* 95-171.

Vemuri, S. and Bender, W. (2004). Next-generation personal memory aids. *BT Technology Journal, 22(4),* 125–138.

Vemuri, S., Schmandt, C., Bender, W. (2006). iRemember: A Personal Long-term Memory Prosthesis. In *Proc. of CARPE 2006,* 65-74.

Wheeler, L. and Rois, H.T. (1991). Self-Recording of Everyday Life Events: Origins, Types, and Uses. *Journal of Personality, 59(3),* 339–355.

Yang, S., Congleton, B., Luc, G., Pérez-QuiÑones, M. A., Fox, E. A. (2006). Demonstrating the use of a SenseCam in two domains. *In Proc. of JCDL 2006,* 376-376.

Yeh, R., Liao, C., Klemmer, S., Guimbretière, F., Lee, B., Kakaradov, B., Stamberger, J. and Paepcke, A. (2006). ButterflyNet: a mobile capture and access system for field biology research. In *Proc. of CHI 2006,* 571-580.

FOOTNOTES

(Make a copy of all footnotes on a separate page here. This only has to be done for the final submission for production. During the review process, it is okay to just have footnotes at the bottom of pages.)

1. SenseCam has recently been released as a commercial product, known as ViconRevue. However, when this work was conducted, SenseCam was still available only as a research tool, and thus we reference SenseCam not ViconRevue throughout this work.

2. http://www.nsf.gov/news/news_summ.jsp?cntn_id=116928, http://tochi.acm.org/authors.shtml

FIGURE CAPTIONS

Figure 1. An example of the paratyping survey used in the SenseCam study. Proxies complete the left portion of the survey to documents eligible encounters. The right portion of the survey is a pre-stamped postcard which contains questions to be answered and mailed by survey recipients as well as a description of the SenseCam system.

Figure 2. Key features of the paratyping method, and the advantages and disadvantages inherent to each of those features

Figure 3. (left) Personal Audio Loop system in its prototype form used for pilot deployment study. (right) The review interface is largely audio-based, but there are some visual elements to help users recognize their mode and where they are in the audio buffer when playing back sound.

Figure 4. (left) SenseCam and (right) the interface used to review images. Salient moments can be flagged with bookmarks, and the images can be reviewed at various speeds.

FIGURES

Figure 1. An example of the paratyping survey used in the SenseCam study. Proxies complete the left portion of the survey to documents eligible encounters. The right portion of the survey is a pre-stamped postcard which contains questions to be answered and mailed by survey recipients as well as a description of the SenseCam system.

Date: _____ / _____ / _____

1) What were you doing / talking about?

2) Sensitive information involved:

None Financial Medical Other

3) Physical location:

4) Describe the people around camera's field of view

5) How did the recipient respond to you giving them the survey?

6) Notes (include your relationship with the person but NOT the person's name)

7) If you chose not to hand out the survey, please note that here and explain

0713

SenseCam is a wearable digital camera that is designed to take photographs automatically, without user intervention, while it is being worn. SenseCam takes small low-resolution pictures and stores them on internal flash memory. Unlike a regular digital camera or a cameraphone, SenseCam does not have a viewfinder or display. It has a wide-angle (fish-eye) lens that maximizes its field-of-view. The device doe not capture sound. SenseCam can operate on a timer, for example taking a picture every thirty seconds. Also, a big change in light or heat levels around the camera can make it take a picture. Once transferred to a computer, people can watch the images slowly, quickly, re-wound or paused. They can also delete individual images.

Other researchers are looking at how the SenseCam might help people as a memory aid. **Our team wants to know about your feelings about it as a person who might have been recorded.** There are no wrong answers.

If you are willing to join this research, suppose that the person who gave you this card is using SenseCam. Then, please complete the survey as soon as possible and drop it in any mailbox. This survey should only take you a few minutes to complete, and there are no costs to you but your time. There are also no benefits to you, but we hope that better understanding of people's reactions to SenseCam will help others in the future. The risks of this research are no greater than you might encounter in daily life.

The person who handed you this card will never see your answers. Our research team will keep your answers anonymous and confidential. By returning your completed survey, you are agreeing to join this research project. You do not have to return this survey. This person who gave it to you will never know if you returned it or not. **Thank you!**

1) How important would it be that she had told you at the beginning of your encounter that SenseCam is running?	Does not matter	1	2	3	4	5	Matters very much
2) How important would it be that she had asked for your permission to use SenseCam?	Not important	1	2	3	4	5	Very important
3) How likely would it be that you ask her to erase the recording of the encounter you just had?	Not likely	1	2	3	4	5	Very likely
4) How important is it that she asks for your permission to play the recorded encounter to someone else?	Not important	1	2	3	4	5	Very important
5) Do you consider this encounter confidential?	Not confidential	1	2	3	4	5	Very confidential

6) What was the nature of your encounter?
☐ Work ☐ Personal ☐ Gossip ☐ Logistics ☐ Transactional ☐ Other _____

7) Your Age Range: ☐ 18-29 ☐ 30's ☐ 40's ☐ 50's ☐ 60 or over

8) Your Gender: ☐ M ☐ F

9) Your Occupation: _____

10) Today's date: _____ / _____ / _____

11) If you want to participate in a 1 hour follow-up interview with $20 compensation, please write an email address or phone number below. Or if you are more comfortable, email us at gillianrh@ics.uci.edu to volunteer.

turn card

0713 _____

Figure 2: Key features of the paratyping method, and the advantages and disadvantages inherent to each of those features

Feature	Advantages	Disadvantages
Engagement with proxies	• Creation of an authentic instance of use • High number of survey recipients that scales with proxy scale • No need for fully functional prototypes • Less risk from use of prototype systems, fully functional or not	• Training is required to reduce bias • Bias still exists in sampling, self-report from both respondents, and proxies, and so on • Ethical questions surrounding the role of the proxy in the research and the ability for survey respondents to gather more information about the study
Limiting survey question length and quantity	• High response rates • Responses close in time to the encounter	• Limited understanding of underlying rationale for responses • Questionnaire wording must be precise
Standardized system descriptions	• Improved reliability of information delivered to survey respondents • Reduction of proxy bias influencing survey responses	• Limited information that can be conveyed • Descriptions must be carefully validated prior to deployment of the instrument
Anonymous surveys	• Reduced risk for survey respondents • Potentially increased honesty from survey respondents	• Difficulty contextualizing responses • Potential for fraudulent behavior on part of proxies
Follow-up interviews with survey respondents	• More depth in understanding of participant responses	• Identification of survey respondent, which can lead to bias in response

Figure 3. (left) Personal Audio Loop system in its prototype form used for pilot deployment study. (right) The review interface is largely audio-based, but there are some visual elements to help users recognize their mode and where they are in the audio buffer when playing back sound.

Figure 4. (left) SenseCam and (right) the interface used to review images. Salient moments can be flagged with bookmarks, and the images can be reviewed at various speeds.