TagPad for iPad – Designing a Support Tool for Interview Studies

Nis Bornoe\textsuperscript{1}, Louise Barkhuus\textsuperscript{1}, Barry Brown\textsuperscript{2}, Malcolm Hall\textsuperscript{1}

\textsuperscript{1}Department of Computer Science
University of California, San Diego
La Jolla, CA, USA
\{nis, barkhuus, mhall\}@cs.ucsd.edu

\textsuperscript{2}Department of Communication
University of California, San Diego
La Jolla, CA, USA
barry@ucsd.edu

ABSTRACT
In this paper we present the iPad app TagPad. The app is designed to support interview studies. It's been designed to fit flexibly with structured and semi-structured interviews for both small and large-scale studies. TagPad can record audio and save text entries and multiple-choice answers. The data can be uploaded to a cloud-based data repository making it simple to access and share data. To get a better understanding of how interview studies are conducted, particularly the data collection phase, we looked into 12 ongoing studies and found that the data collection often is associated with unorganized and complicated processes. We present the results and briefly discuss how TagPad can optimize and support these processes.

INTRODUCTION
We here present and provide an overview of TagPad – an iPad app designed for conducting interview studies. It's intended to support the data collection phase and also offers the possibility to add tags to the interviews for pre-analysis. The idea is to provide a tool that can support research and automat some processes while still offering the researcher flexibility and creativity so it's not the tool dictating the data collection. As Barbour points out qualitative research should not be reduced to a list of technical procedures. It should not be: “the tail wagging the dog.” [1].

Not only are we interested in developing a usefulness tool. We also wanted to better understand the data collection phase in interview studies and the context in which studies are being conducted. The data collection phase is an area that still needs more investigation. Such knowledge can help determine how collecting data can be made more effective and reliable. Other research has shown that when developing tools for researchers the settings and requirements need to be understood [4]. By examining ongoing studies we investigated how data collections are carried out and in this paper we present the results. Running trials only to determine if an application can be considered a success or not is neither that useful or interesting. It’s both difficult to decide when a trial can be labeled as a success and to get a broader understanding of the system and the diversity of the users [2]. Therefore domain knowledge is part of the development process. Not only is this knowledge useful when designing features and interface. The knowledge will also be useful for further studies of TagPad including setting up field trials and to better understand the feedback gained during these trials. When designing and evaluating a research tool it’s also about looking into how the tool positively can change existing procedures, deciding on good practices for using the tool, and not least looking into the limitations of the tool. The intention with this paper is to introduce TagPad and provide a summary of the information gained from examining 12 ongoing studies.

THE TAGPAD APP
TagPad\footnote{1} is an iPad app designed for researchers conducting interview studies. The app can be used for qualitative studies or quantitative data collections such as in-person surveys and is designed both for in-person and phone interviews. It's been designed to fit as flexibly as possible with structured and semi-structured interviews. The basic concept is that an interview guide/instrument is loaded into TagPad (See Figure 1.) and different input can be added to the different questions. As input TagPad can record audio, save short text entries and multiple-choice selections. The researcher decides how to combine the input. For example, TagPad can purely be used to record audio, it can be used for only short text entries and/or multiple-choice questions or a combination, so some or all answers can be recorded while some or all questions can be answered with short text entries and/or multiple-choice options. Because of this flexibility TagPad can be considered a multi purpose interview tool suitable for a wide selection of qualitative and quantitative studies. The output generated by TagPad consists of an audio file containing the entire interview and individual audio files for each question. For flexibility audio files are saved in the MP4 format (.mva) recognized by common audio players. Text entries and multiple-choice options are saved in a comma-separated values (CSV) file.

\footnote{1} The version described is 1.02. Latest information about TagPad can be found at http://www.tagpad.info
A vital feature of TagPad is the platform it’s running on – the iPad. This tablet computer is partly defining TagPad because it offers the interface and input sensors needed all in a lightweight portable device. This platform was found suitable for the circumstances and environments many studies are conducted under. For example, an iPad can be operated while standing up and is convenient for interview sessions or observations requiring mobility. In comparison to a laptop the iPad is less of a barrier and less intruding during social interaction. In addition the iPad can agilely be prepared for an interview and operates fairly long on battery without the need for an external power source. With TagPad it was an aim to design a “turnkey” app that easily can be integrated into many different projects and to design an app that does not require a high level of technical expertise to operate. To achieve this we focused on developing an app with a minimal set of features and options and an app producing output in common formats.

EXAMINING THE DATA COLLECTION PHASE

We looked into 12 different ongoing studies to get a better understanding of the data collection phase in interview studies. All studies were either in the study design or data collection phase. A representative for each study explained us the study outline, time frame, and which participants they were looking for. We asked how they would organize the data collection, which methods they would use, by what means they would collect the data and which questions they would ask. To get an authentic picture we requested answers based on actual experiences and not how they ideally wanted to or originally had planed to collect data but how it actually occurred. We looked into both small and large-scale studies, from research teams consisting of only one person conducting 10 – 20 interviews to research teams of 15+ planning to conduct up to 1000+ interviews. Out of the 12 studies eight can be classified as small-scale studies with one to four team members completing 10 – 25 interviews. Four are considered large-scale studies with five+ team members and 100+ interviews already completed or planed. The studies conducted qualitative interviews, quantitative surveys and observations. A mixed approach was common, for example, in one study observations would be mixed with interviews.

A Closer Look at the Interview Studies

In the small-scale studies the involved researchers would complete all stages, from designing the study, to recruit participants, write the interview guide, conduct the interviews, analyze the data and write a rapport or paper. Occasionally some subtasks such as transcription would be outsourced. In the large-scale studies the research team would consist of several sub-teams. These teams can in general be divided into an executive group that would be in charge of the study and involved in the analysis, a group would collect the data, and a group would be involved in analyzing data chunks. Often there would be overlaps between the different groups, so some data collectors would also be involved in the analysis process etc. To illustrate we

---

**Figure 1:** An interview guide is loaded into TagPad. Here question 5.2 is selected. It is possible to audio record the answer and/or type in a short text entry.

**Figure 2:** This is the analysis view of TagPad. Here it’s possible to add custom tags to each interview for pre-analysis.

TagPad has a simple analysis view for quick analysis to support selective transcription etc. (See Figure 2.). Besides having playback capabilities tags are automatically added to the audio timeline for each question so it’s possible to locate where in the audio a specific question was asked. Custom tags such as “Good quote” can be added making it easier to navigate, locate and share specific audio segments. Tagging is possible both during and after an interview.

TagPad uses cloud computing for data storage and currently uses the service Dropbox [3]. This file storage service synchronizes files between connected devices such as, desktops, laptops, tablets, and a Web accessible file repository. Dropbox was chosen because it’s free, widely used and has some built-in collaborative features including easy file and folder sharing and auto backup. Despite using a cloud-based data repository, Internet access is not needed during interview sessions, as all data also is stored locally.
here present two studies with very different proportions. The first study presented is a joint project between eight American and Mexican universities. In the second study presented the research team consisted of a single person – a grad student collecting data for her master’s thesis.

**Large-Scale Study: Mexican Deportees Survey Study**

In this project deported Mexican immigrants are asked to participate in a survey regarding their illegal entrance into the US and their deportation back to Mexico. The aim is to get a better understanding about what happens to Mexican deportees. Potentially politicians, think tanks, and activist groups can use this information.

The participants are mainly recruited and interviewed at shelters and immigrant drop-off locations such as bus stations at the border. They are in the age range 18 - ~50 years old and the average age is about 32. Often the participants are in an intermediate stage, using shelters for a short period of time before they will try to reenter the US or move on to other locations. A large majority are Mexican citizens, and the remaining subjects are citizens of other Central and Southern American countries. The goal is to conduct 1000 – 1500 in-person interviews. The paper-based survey consists of about 250 questions, and typically about 170 questions apply to a given participant and takes about 45 minutes to complete. The survey consists of a mix of open-ended and multiple-choice questions. For the open-ended question, answers are written down in one or two short sentences or keywords. Rarely interviews are audio recorded. After completion the paper-based surveys are scanned and forwarded to a team manually typing in the data from the scanned surveys into a spreadsheet. The quantitative data will be statistically analyzed and it’s still not clear how the qualitative data will be analyzed. The study is a joint project between eight American and Mexican universities and overall 15+ people are involved in the project. Researchers and in particular grad students are participating for shorter periods and typically grad students will collect data as part of their studies. An executive group is designing and leading the study. The group will in collaboration analyze the data and plan to meet for a series of workshops to coordinate the analysis. It’s the plan to repeat the study in a few years for comparison purposes.

**Small-Scale Study: Study of Food Habits Among Pacific Islanders Living in San Diego**

This study explores the role of traditional foods among the San Diego Pacific Islander community. The project looks at connections between food, health, and identity. An applied anthropologist student using the data for her master’s thesis conducted the study. She is the only person involved in the project and will complete all steps of the study from study design to data analysis and final write up. The aim is to complete 20 interviews. All participants are pacific islanders living in San Diego and are in the age range of ~20 - ~80 years old. Participants are participating in a single session lasting about 45 minutes. Interviews are taking place at different public locations such as cafés and at the beach. The interview guide consists of very open-ended questions and often people will start talking about anything under the sun. Occasionally the researcher will e-mail the participants afterwards for clarifications. The interviews are being audio recorded and she uses a paper-based interview guide. Notes are taken on a piece of paper, for example, spelling of unfamiliar food names etc.

**HIGHLIGHTS FROM THE INTERVIEWS**

The use of paper-based interview guides/instruments was dominating. Notes were taken on notepaper or laptops (or both) and/or noted in the margin of the interview guide. Researchers recording audio used simple external digital audio recorders. In general recording of audio only applied to the small-scale studies. The interview guide was in most cases very dynamic and would be changed over time, as irrelevant questions would be removed and interesting topics further explored. Using a paper-based interview guide allowed rapid changes during the interviews.

For a large majority of the interviews the researchers would commute to meet with the participants. This is often the case because of practical reasons, for example, when interviewing patients at a hospital, and when the nature of a study requires it such as when doing field observations. This also means that access to Internet and an external power source is not guaranteed.

We found that the processing of the collected data often is complicated, unorganized, and involves several links. In the large-scale studies involving hundreds of in-person interviews or surveys a common approach was to scan the completed paper surveys and then forward the digitalized versions to a team manually processing the data. This process involved several people and is time consuming. Similarly in the small-scale studies processing was done manually including uploading audio files and notes to a computer, organizing the files, and creating a backup. Notes would both be saved and organized in text files and on paper. Particularly the management of notes was very different from study to study.

Simple tools were used for managing the data. For example, word processing software was often used both during the data collecting and the analysis. The use of dedicated Computer Assisted/Aided Qualitative Data Analysis software (CAQDAS) was uncommon. Particularly in the small-scale studies it was common to use whatever software was available and possible to get for free or low cost. Using technology to support the data collection was something our participants was interested in but they had trouble finding applicable software.

Securing collected data was an explicit concern because of the resources needed to collect data and because an observation or interview can be hard to do over again. This was a central reason for using low-tech tools such as pen and paper because they are considered reliable. High-tech solutions were found more risky and stressful to use. When recording audio the quality was often questionable. Often
the recordings would contain a lot of background noise. This was particularly an issue when doing phone interviews because the audio was recorded using the phone’s loudspeaker resulting in heavily reduced audio quality.

Audio recordings were almost exclusively being transcribed and then not used again during the analysis. It’s not uncommon that audio recordings can last an hour or more so navigating and effectively finding the right segments has been pointed out as challenging. Instead the transcription is used to single out interesting and relevant segments.

We found that the strategies for both the data collection and analysis were only prepared on a very general scale. This applied to all aspects of a study from recruiting participants, deciding on interview locations and interview strategies. Mainly a general study design was decided beforehand, all remaining parts would be dealt with ad-hoc.

DISCUSSION
Based on our examination we identified two central elements of the data collection phase. One element is the actual collection when a researcher conducts an interview or observation and saves data such as field notes, audio, video etc. Interview studies are conducted using a paper-based interview guide, notepaper or a laptop for field notes, and an audio recorder. The idea with TagPad is to create a tool that combines these tools into one integrated solution while utilizing the features of the iPad. The iPad with TagPad installed provides a powerful IT instrument that can operate without Internet access, has a relative long battery life and has the option to connect to the mobile Internet. As mentioned the format of the iPad itself is an essential factor because it offers the interface and input sensors needed all in a lightweight portable device with a screen size still large enough to be useful. This applies to several levels. First off the size and lightweight is important because researches often commute to meet participants and to conduct field observations. Secondly an iPad blends in well with most settings. For example, recording audio during observations can feel less intrusive to the participants because it’s not sensed as an explicit audio recorder. During interviews the iPad can be less of a barrier in comparison to a laptop.

Another element of the data collection phase is the processing and management of the data. This includes organizing, storing, archiving, and preparing the data for analysis. This step has several subtasks. For example, digitalizing and transferring the data to a spreadsheet, database, word processor or CAQDAS. This part of the data collection was found to be challenging and time consuming. By automating some of these tasks, such as a simplifying data uploading, a goal with TagPad is to make this part less challenging. Combining TagPad with a cloud-based data repository adds further advantages including the support of auto backup and more options for data sharing and collaboration. Based on our examination we believe both small and large-scale studies can benefit from these advantages. In particular we believe TagPad can support scientific collaboration because the data rapidly can be made available online in common data formats and this can be done with minimal technical skills. In large-scale studies having several people collecting data, TagPad can help both streamline the data collection making it more consistent and support a smoother and less complicated uploading process. Because conducting a study is time consuming and complex and often involving people with different backgrounds and skills we have focused on designing a tool that is fast and easy to learn. Our approach was to implement a minimal set of features and automat some processes. It was clear that our participants are anxious to loose data so feeling comfortable with the tools used is absolutely essential.

We also recognize some potential flaws. TagPad can be too inflexible for certain studies. Especially interview guides based on very open-ended questions because the interview guide cannot be modified during an interview. Another issue is that tablet computers are not suitable for all environments, for example, some outdoor conditions such as rain and direct sunlight. Further a tablet computer might not fit into all interview settings and can potentially receive too much attention and focus and act as a barrier between the interviewer and interviewee.

CONCLUSION
We have presented the iPad app TagPad designed to support interview studies and in-person surveys. We have also looked into the data collection phase in both small and large-scale studies. Our goal was to introduce the features and design of TagPad and provide a better understanding of how data collections are conducted in reality. This is both useful for further development of TagPad and to consider how research can be made more reliable and effective. We found that a major challenge was processing the collected data and this was done unstructured and ineffective. With TagPad we believe it will be possible to better streamline the data collection and automate some processes such as uploading data.

ACKNOWLEDGMENTS
This study was funded by the National Science Foundation (NSF) grant #0838330.

REFERENCES
3. Dropbox
http://www.dropbox.com