Rapid Prototyping and Field Evaluation of Mobile Experiences

Frank Bentley

Experiences Research

Motorola Applied Research and Technology Center.

1295 E. Algonquin Rd.

Schaumburg, IL 60196 USA

f.bentley@motorola.com

Crysta Metcalf

Experiences Research

Motorola Applied Research and Technology Center

1295 E. Algonquin Rd.

Schaumburg, IL 60196 USA

crysta.metcalf@motorola.com

Abstract

We present a set of guidelines for mobile experiences research that we have evolved over six years of creating and testing mobile concepts at Motorola Labs. We focus on guidelines for deciding what to build and how to run field evaluations.

Keywords

Concept Testing, Ecological Validation, Rapid Prototyping

ACM Classification Keywords

H5.m. Information interfaces and presentation (e.g., HCI): Miscellaneous.

Introduction

When conducting exploratory ethnographic-style research, we often complete each study with dozens of design ideas for new applications and services to create based in the data that we have collected. Through initial exploratory studies, we seek to understand the potential areas where new services and applications can have the most impact. From this data we move forward into building prototypes of, and testing the various concepts embodied in the prototypes. We usually have two main goals for building a prototype and conducting a field concept study. First, we use this work as risk-mitigation for continuing research in a particular area. By quickly testing a concept we can

see if there is any benefit to the proposed experience. Second, we often have specific research questions about a type of experience or interaction and we use the prototype to get valid feedback from users in real contexts of use.

Through the course of our work exploring Ambient Mobile Communications [3], we've developed methods for building and evaluating these prototypes in the field that we believe are useful for other mobile researchers. Because mobile interactions are very context dependent, we strongly believe that testing in the field is critical to understanding how a particular concept will be used in daily life.

What we test

We have two main principles that guide us when implementing a concept to test in the field. These principles help us quickly create a functional application that a participant can use over the course of several weeks in their daily lives. These principles are to build only what you need and to build the experience, not the technology. Through following these guidelines, we ensure that we spend the least amount of time on a given project until we prove out the key components of the experience. If the idea does not turn out to be a good one, we have wasted little time on it.

Build only what you need

Because of our goals, our prototypes tend to be made rapidly, often in a matter of a week or two, and are built to test a specific concept with users. We prototype the key use cases and leave the less-relevant functionality out. For example, our Motion Presence [2] application replicated a mobile phone book, but didn't bother creating ways to add new contacts or add all of the rich fields (e.g. birthdays, addresses, etc.) that are present in a full-featured phonebook. Rather, we displayed the name, number, and motion status (the concept we were testing) and provided options to call or text message that person. In this way we could rapidly test the concept without building out a complete application.

Likewise, in the Music Presence study [1], we provided people with knowledge of the music that their friends were playing in order to see how participants would use this information in their daily lives. For this study we didn't even build a mobile application, we just enabled a server to send text messages about friend's music listening straight to our user's devices. By keeping the interaction simple, we could quickly test the potential of this information; it took just a few days work. While receiving this information as an SMS is not the ideal user experience, we could learn how our participants used this data in their daily lives to coordinate, start conversations, and send subtle messages to others through music choice. This showed us that there was promise in the concept, and we moved on to higher fidelity prototypes. We could not have learned this information through paper prototyping or lab studies, it was critical to build a functional experience and test it in the world with real participants.

Build the experience, not the technology

Because we are interested in the experiences and concepts, the prototypes that we build are often not engineered in the way that a commercial offering would be implemented. Because we utilize small-scale studies, with limited numbers of participants, we have the ability to choose the most expedient path for implementation and have the ability to subsidize data

or SMS plans for our participants. In Motion Presence, we simply sent SMS messages to a particular port on the phone any time the motion status of a friend changed. This could be implemented quickly and because we were using small groups, it worked. A commercial solution of this application might tie into complex IMS/SIP presence systems and would have taken much longer to implement. We believe that testing the overall concept is the most important first step. If the concept fails there's often no reason to spend the time to develop the technology further.

However, our concepts need to be built sturdy enough to survive everyday use for multiple weeks in the field. By building simply, this makes it much easier to reach the point where significant additional work would not make the application significantly more robust. In times when we relied on working with advanced technology components or fully developing an interface, we often found the experience to be lost in the complexity or the new technology did not perform to our needs for the given experience. When this happened, we went over our anticipated deadlines, and in one case, we had to stop a study entirely because of the state of the underlying technology. We would much rather test the concept using alternative implementations and fully develop the appropriate commercially viable technology once the concept is more fully proven out.

These guidelines have allowed us to be quite rapid in discovering key experiences that work in the lives of our users. Often, when we finish an exploratory study, we have hundreds of design ideas to choose from. Taking a few of the most promising and quickly prototyping them and testing them in the lives of users helps us to identify which are the most promising to pursue for further research or for commercialization.

How we test

Once we have the prototype created, we have a number of different methods that we use to evaluate the experience. While these methods vary based on what we are testing and our research questions, the choice of our combination of methods is always based on a single guiding principle: we want the users to engage with the prototype in the course of their daily lives. Ecological validity, to us, does not mean approximating the real-life situation; we want the materials and settings to be the real-life situation. There are so many contextual variables that affect the mobile experience – where the person is located, who else is around, their current needs and emotional states, just to name a few. Putting a prototype into users' hands for a week to two weeks gives us a good indication of how people will actually use and react to the concept. Testing in a lab, or with paper prototypes will not get us this rich interaction information.

Because of this guiding principle, in our research we do four things that we consider the basic criteria for how we test our prototypes. First, we recruit social groups (people who already know each other) when we are testing social technologies, versus asking strangers to act as if they know each other. Second, we put the technology in the field: we ask people to take it home with them and use it as they would if they were not in a study. Third, we make sure that the participant only needs to carry one mobile device. And fourth, we select data collection techniques that allow us to come as close as possible to "being there".

Social groups for social technologies
Many of the applications we have worked on have been social applications – applications that were created to enhance experiences for people who already know each other. Because of this, we think it is essential to recruit participants who are friends or family members. It is the knowledge of, and interest in, the other people's lives that makes the experience meaningful. For example, in our motion presence study it was knowledge of her husband's schedule that allowed one of our participants to interpret that "moving" meant her husband was not yet in his meeting. If strangers are recruited to evaluate technology that is fundamentally social in nature, not only will they not be motivated to be social with the other participants, they will not have the background knowledge they need in order to be fully involved in the others' lives. This diminishes their ability to evaluate the experience.

Real context of use
When we put prototypes in the field we do so for at least one week, and in most cases we have it in the field for two to three weeks. While we ask our participants to use the application or feature "enough to evaluate it," we do not ask them to use certain features or functions any number of times, on certain occasions, or for any particular reason. Indeed, we tell them that they should use it as if they are not in a study, using it only as they see fit. Part of our evaluation is discovering when they think it is appropriate to use the prototype (or not appropriate to use it as the case may be), and what they are interested in using it for.

For example, in the Photo Presence study [1], we saw participants using the application to share images of inside jokes or just typical imagery of what they saw

and experienced throughout the day. These real contexts of use allowed us to see how photo presence as a concept fit into people's lives.

Primary device
We also believe that in mobile studies of experiences, making the participant carry an extra device seriously detracts from the ecological validity of the research. When the final solution is envisioned as something that would exist on a primary the device, we recognize that use will be different on a second device that is not looked at frequently throughout the day for other purposes. Furthermore, the participant will be less likely to carry around and use the second device, especially if the only thing it does is the thing we are studying. In sum, the participants will be less likely to engage with the prototype the way they would in real life. In order to maintain our one device principle we transfer the participants' SIM cards (along with their contacts and all other information they need), to the prototype device. It makes our job more difficult, but makes engaging in the study less work for the participant, and again, more closely simulates a real-life experience.

Field-based data collection
While we use semi-structured interviewing in all of our studies in order to delve into the reasons for use (or lack thereof), we also combine this technique with supplemental methods that get us closer to "being there." Logging key presses with time-stamps, for example, lets us know exact usage patterns during the study. We also use voice mail diaries – so participants record their thoughts as close to the time of their experience as possible. In some cases we have recorded our participants' phone calls (of course with

their permission and the permission of the other parties on the call). The use of these techniques is based on the premise that we collect better data the closer we get to direct observation of real-life use. While experience sampling is another excellent technique, it often induces interactions at unrealistic times and does not capture data at the time of actual desired use. Our experience with voicemail diaries has given us the information we need, relatively close to the time of actual user-driven interaction. We believe that these techniques are appropriate and essential ways of collecting more accurate data than relying solely on interviews sometimes days after particular interactions have occurred.

Conclusion

We hope that these principles can help guide other researchers in validating concepts with real participants in the context of everyday use. We believe that these types of studies out in the world are critical in rapidly understanding the potential uses of a particular application or service offering and in evaluating the viability of experiences before investing a great deal of effort on implementation.

Acknowledgements

References

[1] Bentley, F. and Metcalf, C. The Use of Mobile Social Presence. IEEE Pervasive Computing. 2009. (to appear)

[2] Bentley, F. and Metcalf, C. Sharing Motion Information with Close Family and Friends. Proceedings of the SIGCHI conference on Human Factors in computing systems (CHI 2007), San Jose, CA, USA, 2007.

[3] Bentley, F., Kaushik, P., Narasimhan, N., and Dhiraj, A. Ambient Mobile Communications. CHI 2006 Workshop on Mobile Social Software, Montreal, Canada, 2006.