Evaluating Early Prototypes in Context: Trade-offs, Challenges, and Successes

Early-prototyping approaches that ultimately sacrifice some realism or allow some bias in order to approximate contextual evaluation can provide useful evaluations of design choices.

Derek Reilly, David Dearman, Michael Welsman-Dinelle, and Kori Inkpen
Dalhousie University

Pervasive computing's mixed success as a paradigm in everyday life makes evident the need for evaluation methods that provide insights into likely usage patterns.[1] Focused research questions tend to be explored in lab experiments,[2] while larger projects often involve implementation and field evaluation of a completely realized concept. It's in the interest of designers and researchers alike to achieve the middle ground: that is, contextual evaluation of early or incomplete prototypes. However, getting participant buy-in can be difficult, results can be misattributed owing to prototype deficiencies, and a host of other methodological challenges exist. We've investigated two approaches that let us evaluate early prototypes by achieving useful approximations of evaluation in context. We used these approaches to evaluate prototypes designed to facilitate navigation in groups.

The case for approximation

Pervasive computing applications aren't straightforward to create. The up-front costs of hardware and software infrastructure are usually very high owing to a lack of standards and know-how and to a need to use expensive, emerging, or experimental technology. Users often must learn new skills for interaction and alter their activities in ways that are perhaps more disruptive than those presented by "traditional" desktop applications. It's unsurprising that so much research in the area has occurred in controlled laboratory settings or that development sometimes bypasses prototype evaluation altogether.

In addition, ongoing evaluation of prototypes is important. A cautionary example is the mobile guide developed for the Exploratorium, an interactive science museum in San Francisco.[3] Design of a complete working prototype was based on a needs analysis for museum visitors. The developers then evaluated the resulting prototype on site, only to find that the mobility and hands-on nature of the exhibits made the guide cumbersome to use most of the time. A complete redesign led to a passive solution that implicitly collected experiences for the museum visitor to retrieve later. Such late revelations are costly.

Ideally, developers would evaluate prototypes in context earlier and continually during development to avoid the need to scrap fully articulated designs. However, without a robust, functioning system, getting the buy-in required to permit true contextual evaluation is difficult.[4] This is particularly pronounced in pervasive applications, which are integrated with the environment and context of use.

Our two approaches address this seeming

Published by the IEEE CS and IEEE ComSoc ■ 1536-1268/05/$20.00

conundrum by recognizing the challenges of evaluating early prototypes in context, accepting certain limitations yet providing real evaluative power. The first, a variant of *experience prototyping*,[5] permits some bias in evaluation to allow use of prototypes in real situations. The second approach, which employs *Wizard of Oz prototyping* or *functional approximation*, sacrifices some realism to evaluate the prototype with impartial participants. By acknowledging and working with these approaches' inherent limitations, we've been able to evaluate a range of early prototypes. The results must be viewed in relation to the limitations but can provide valuable insight into the suitability and effectiveness of designs for pervasive computing applications.

Group navigation

Technology supporting group navigation might increase the quality of communication by promoting awareness of one another's actions and intentions or by providing shared visual aids. It might decrease the chances of navigation error by making information available when needed or by supporting group problem-solving approaches. Motivated by our own experiences traveling in groups, we identified three lightweight "technological interventions" and explored each in isolation, allowing distinct lines of inquiry to develop as appropriate. In combination, the research helped us better understand how technology might support group navigation.

Coordinated Views, the first technological intervention, considers how sharing common views and annotations on handheld computers might assist groups navigating together in close proximity. For example, when viewing a map that's larger than the handheld screen, a group can automatically converge on the same portion of the map. Sharing annotations might let the group better communicate and record plans.

Marked-Up Maps looks at how to use RFID-tagged paper maps with handheld devices to support navigation and information retrieval. The impetus for this research was to consider the common situation of several people sharing one paper map. Adding handheld interactivity might permit capturing chunks of the map for collaborative search tasks, while maintaining a large, shared view (the paper map).

Rendezvous explores different technologies' impact on the act of meeting at a negotiated time and place. Technologies studied include a location-aware handheld application and regular mobile phones. Scenarios include establishing a meeting place dynamically, responding to lateness, and making last-minute changes to meeting plans.

Partly because we're most interested in understanding the problem domain and not explicitly trying to develop a complete system, the need for contextual evaluation has driven our prototyping efforts from the outset. However, our experiences also apply to those who are more interested in designing real applications.

Experience prototyping

This approach emphasizes "the experiential aspect of whatever representations are needed to successfully (re)live or convey an experience with a product, space or system."[5] As an overall philosophy for prototyping, you can apply it to understanding existing practices, communicating design ideas, and evaluating the prototype. One concrete approach to experience prototyping is for designers themselves to become immersed in the target population's current practices to better understand their needs and perspective.

One concrete approach to experience prototyping is for designers themselves to become immersed in the target population's current practices.

In our research, we've applied an experience-prototyping technique in which researchers are participants in actual scenarios; that is, they're not just pretending to be members of the target population. Researchers or designers mediate their own authentic experiences using prototypes in context, yielding powerful, visceral impressions that can provide insight when subject to careful reflection. The participating researchers' desire for the activity to succeed can mitigate their bias toward the technology (we discuss this bias in more detail later). Designers can also more effectively manage the early prototypes' rough edges; there's less possibility of misinterpreting a problem with a prototype as a problem with the design. Finally, directly confronting the design assumptions' impact is invaluable to researchers and designers when they're refining their prototype.

To investigate experience prototyping, we used Coordinated Views and Marked-Up Maps.

Coordinated Views

We explored the use of our Coordinated Views software during City Chase (www.thecitychase.com), an organized city scavenger hunt. Three pairs of researchers participated in the race. During the race, each pair had access to our Coordinated Views prototype.

The prototype provided shared views and annotations using Bluetooth peer-

(a) (b) (c)

Figure 1. Participation in a city scavenger hunt: (a) Coordinated Views software on paired iPaq handhelds, (b) a team planning its next move, without the handhelds, and (c) the pair on the move, with an observer in tow.

to-peer communication on iPaq handheld computers (see figure 1). We identified suitable documents for use in our scavenger hunt: a bus schedule and a transit map with landmarks. We developed the prototype implementation around the use of these two documents and implemented logging to capture interactions. Users could continually synchronize with their partner's view or maintain an independent view. A button toggled the function of the stylus between panning and annotating the document. Annotations appeared on both PDAs.

We decided to permit the use of other resources beyond the Coordinated Views prototype. This would let us continue the race if something went wrong with the prototype or if we had difficulty managing the prototype during more physically demanding points in the race. These resources included direct alternatives to the software, including laminated paper maps and bus schedules, and constant access by cell phone to a "control center" of people ready to research an information request.

The participants had high expectations of the software's utility. This remained true even after a pilot trial in which no one used it. They largely dismissed this as being due to the pilot's condensed course, which was in a very familiar part of town and involved familiar landmarks. However, the software remained largely unused on the race day as well.

We conducted a debriefing after the race. The participants suggested several reasons why they didn't use the software. In addition to the availability of alternate, familiar resources, some participants felt that they were too familiar with the area to require more involved research, which the Coordinated Views software might have facilitated. Certainly the event's time pressure was also a factor in their choices. The only time they used PDAs was when they were tied up in transit. The one pair that actually used shared annotations did so on the ferry, and another pair tried to use the PDA with GPS for positioning only while riding the bus. Finally, weather played a role: it rained heavily during the race's first half, which motivated the participants to use the laminated maps.

Marked-Up Maps

An experience-prototyping opportunity for Marked-Up Maps presented itself in an upcoming conference that several in our group were to attend. Before the conference, we built a prototype that linked a paper tourist map of Nottingham, UK, with basic tourist-related information available on a PDA.

We affixed RFID labels to the back of the map to indicate map locations. We attached an RFID reader to the back of a PDA such that users could query locations by holding the PDA display-side up in front of a map region (see figure 2a). In this first prototype, users retrieved general information about a location by clicking a button on the PDA when it was over the location. Because we couldn't easily connect the reader directly to the PDA, a server on a laptop connected to the reader processed the RFID reads. Pressing a physical button on the PDA (mapped to a request for the server's base URL) caused the server to place HTML-formatted information mapped to the most recently read tag ID at the URL, which the system then displayed on the PDA screen. In this way, the prototype permitted point-and-click interaction with individual locations on the paper map.

We used the prototype while touring downtown Nottingham (see figure 2b). We obtained tourist information from various Web sites about Nottingham and its attractions. Unfortunately, we didn't resolve the prototype's glitches until after the conference, and only one of us remained in the city. That person used the prototype intermittently throughout one day of sightseeing. He took notes as he used the prototype and made summary notes at the end of the day.

This combination of a PDA and paper map, and the wired connection to the laptop in a backpack, made access cumbersome while moving, especially when only one person was using the system. So, the user accessed information from the PDA using the map while sitting

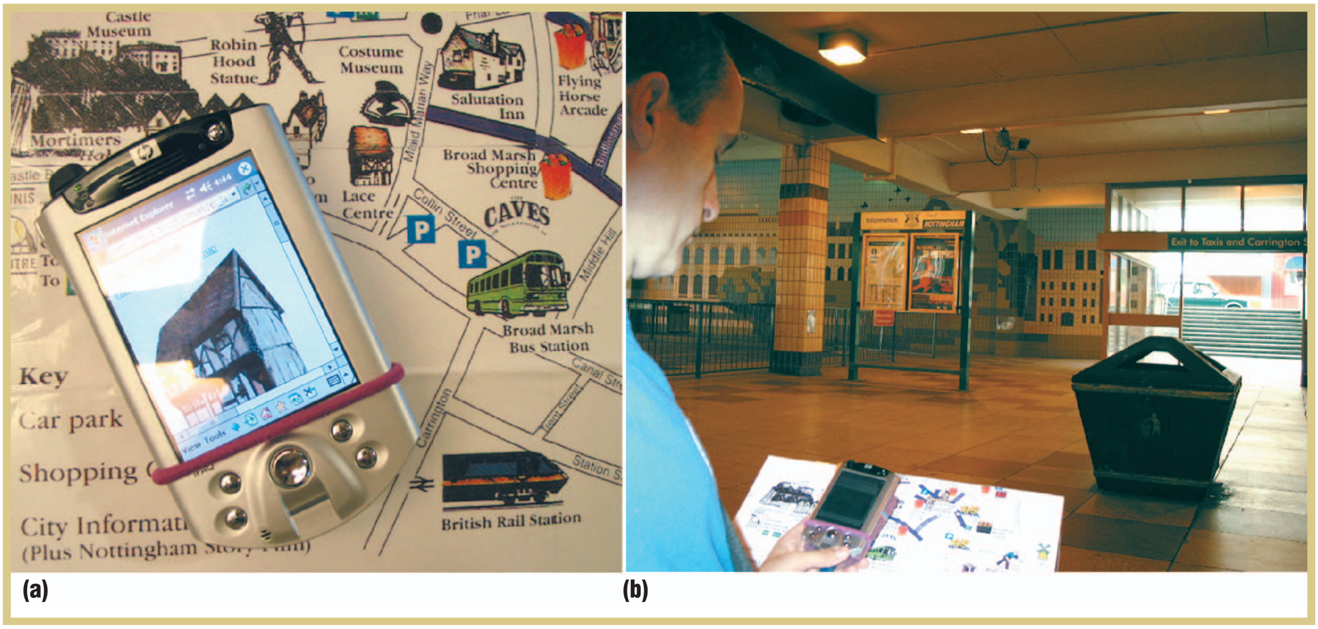


Figure 2. Testing the Marked-Up Maps prototype: (a) Users can obtain details about a tourist attraction by placing a handheld computer over a paper map of Nottingham with RFID tags embedded. (b) A researcher uses the the prototype in Nottingham's Broadmarsh bus station.

down. This was acceptable because he was using the information to plan an itinerary and explore options. Feedback that a read had occurred would have been welcome, however. (Before we used the prototype outdoors we had relied on an audio cue from the server to indicate a read, but this audio signal was too soft to be heard outdoors.)

After a short period of using the prototype, the user determined that framing queries according to themes (for example, admissions and hours of operation, or bus schedules) would have been desirable. However, this wasn't possible with the prototype. Also, the interface was good to browse with, but because the user had manually associated the guide resources (the HTML pages) to the map landmark icons and knew what was "hidden" behind the links, he couldn't really "surf" the map resources. Eventually the PDA became a secondary tool for occasionally retrieving details, and the user kept the map handy while touring.

Lessons and recommendations

In both projects, taking the techniques into an actual usage scenario provided valuable insight. We evaluated our techniques in unforgiving, nonfabricated contexts, giving ourselves as researchers an intimate appreciation for the technique's potential in the context.

Introspection was important during and after each experience. The feeling that we didn't want to use the handhelds during the City Chase was palpable throughout the event. After the event, as a group we could discuss and reflect on our experiences. Our awareness of needs and possible patterns of use for Marked-Up Maps emerged largely on account of the touring scenario. This first-hand experience facilitated the development of subsequent map implementations and a second functional prototype.

The evaluations required fewer overall resources (cameras, observers) than they would had we used nonresearcher participants. We also didn't need to feel that our technology was imposing on the participants' experiences. For the solitary experimenter/tourist in Nottingham, reflection on the application's utility could occur at a natural pace, and tie-ups due to prototype limitations (waking the devices, handling RFID misreads, and missing details in the hypertext) could be accepted without prejudicing the experience as a whole. The City Chase's added time pressure didn't allow for this luxury. We used the Coordinated Views software only when we had enough time and little else to do.

Researchers who are working alone can reflect on their own use in context but might miss some of their own behavior or the impact of some aspect of the technology. When concentrating on use in a real environment, such researchers might forget that they're supposed to also be observing their experience. Even if they're vigilant, they might not be able to detect subtle or characteristic behavior patterns, because these are often more obvious through focused arms-length observation. This might be addressed somewhat if a group conducts the evaluation (as with the City Chase), perhaps including a researcher whose primary purpose is observation, not participation. However, as we experienced in the City Chase, groups can also easily get caught up in the moment, depending on the nature of the activity.

Both these studies involved activities in which we could reasonably participate. They required no specialized knowledge or experience (although the City Chase required a certain level of fitness). Reflective self-evaluation in a real scenario can take place only if the researchers are suitable participants. This quickly becomes problematic in applied domains (for example, archeology and police work). It might be possible in many cases to find a real scenario that shares many aspects of a specific target activity and in which researchers could reasonably take part. Our City Chase scenario fit this description well. This scenario isn't a likely target for technology development but shares enough in common with other group navigation scenarios to be a reasonable test bed or activity in which to explore the Coordinated Views prototype. Additionally, authentic activities for experience prototyping often come with a definite time frame; prototyping needs to manage time carefully to avoid a missed opportunity.

One potential problem would be bias in the researcher's vision of a technique's potential. When researchers are also participants, they need to be particularly careful not to let the interaction experience descend into artifice. It's tempting to get carried away with a new technique's potential or perceived benefits. Expectations of interaction style and environmental impact can cloud your impressions of the actual experience to the point where you might adjust your behavior to accommodate your expectations, no matter how problematic or limiting the device is.

> When researchers are also participants, they need to be particularly careful not to let the interaction experience descend into artifice.

The real-life scenario's impact can be stronger than any researcher bias, however, as was the case during the City Chase. In contrast, the Nottingham touring experience was to some extent driven by the information resource we built to augment the map. A landmark often was of interest because it was linked from the map, not necessarily because it was of interest to a tourist. In addition to researcher bias and expectations for a prototype, researchers are motivated to capture useful information and so might focus unduly on technology use to the detriment of the experience as a whole. This might have limits, however. Even in the Nottingham study, the user used the marked-up map with the PDA only when he was seated in an area that facilitated this use.

We need to be careful when drawing conclusions from personal experiences. As we discussed, several powerful sources of bias and challenges to authenticity are at play when researchers become participants in mobile environments. In the case of the City Chase, it was tempting after the event to dismiss the technique outright and the use of handhelds in such contexts. However, in both cases we gained valuable experiences that aided us during subsequent prototyping.

Wizard of Oz and other reasonable facsimiles

In Wizard of Oz prototyping, a prototype mocks up some or all of the functionality such that it appears to the user to be a functioning system. Functional approximations are often more complete than typical prototypes but sacrifice detail or accuracy in their operation.

To evaluate both types of prototypes, we needed to perform controlled experimental simulation. Both the effort necessary to achieve consistent interaction and required presence of the "wizards" (behind-the-scenes humans who run the simulation) made longer-term or ad hoc use prohibitive for our Wizard of Oz prototype. Our functional-approximation prototype was also incomplete in ways that prohibit its use in real scenarios. For example, the prototype didn't capture gestural interaction with the fidelity that a functioning system would require. The system's response to user interaction was therefore correspondingly coarse grained. While in other domains prototype fidelity might not impact usability results,[6] perceived responsiveness, granularity, and accuracy can significantly affect the user experience of pervasive applications.

Rendezvous

This study aimed to explore how location awareness affects a rendezvous between two individuals. It required a realistic environment to set the stage for the scenarios and immerse the participants in a setting they could relate to and interact with. We had identified a suitable setting in a busy shopping district. However, no suitable location-sensing technology was available in the area to provide positioning data with the fidelity and robustness that the study required. To address this, we implemented a Wizard of Oz prototype to provide and communicate location information.

We intended the prototype to provide the participants with a perceived connection between their handheld computer and a location-awareness service, providing constant updates of their partner's location and some communication facilities. To accomplish this, a wizard fol-

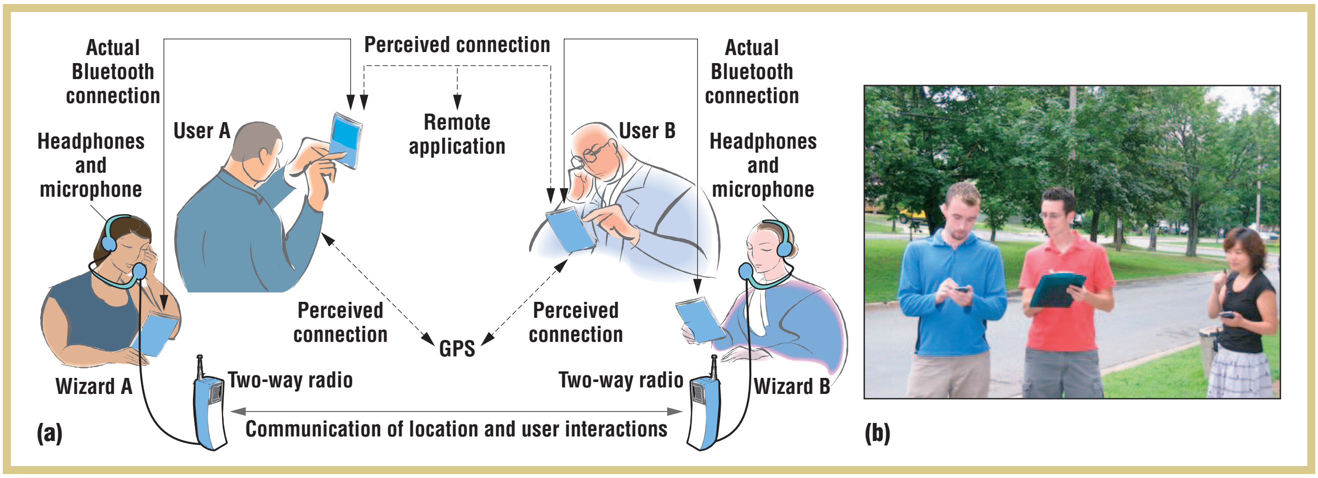


Figure 3. Implementing the Wizard of Oz rendezvous application: (a) The basic approach to location awareness. (b) An observer and a wizard closely followed each participant.

lowed each rendezvousing participant, pushing information updates onto that participant's handheld (see figure 3). Each wizard was responsible for updating the participant's location (using a Bluetooth connection to the participant's handheld) and communicating that location to the other participant's wizard (using two-way radios). During certain scenarios, participants could update the chosen meeting place and request an acknowledgment from their partner. The wizards also facilitated this transaction.

The technique worked well when all proceeded as planned; however, at times the prototype influenced the study. The Bluetooth connection between participant and wizard was subject to range limitations. While we envisioned that the wizard and participant would be in close proximity at all times (see figure 3), this wasn't always the case. For example, several participants would run to make street-crossing signals, resulting in disconnections. In two instances, the number of disconnections became so frequent that we had to throw out the study data. The participants became used to the devices disconnecting. Whenever their partner wasn't shown to be moving, they would ask the wizard if they had lost the connection. In addition, reconnection required a pause in the scenario to let the wizard confirm the connection.

Given the need for proximity, participants sometimes overheard communication between wizards. The shopping district's noise level was often very high, requiring wizards to raise their voice to be heard over vehicle traffic and pedestrian conversations. We observed that this distracted several participants.

We also had to ensure that the participants could take their time to complete the scenarios so that the wizards could keep up while performing their tasks. We therefore gave generous time frames to complete each scenario, where a more restrictive time frame would probably have been more natural.

A final issue related to the participants' perception of the position data's fidelity. When things ran smoothly, the wizards updated positional data with a reliability and precision that intermittent GPS or WiFi positioning couldn't match. So, the user might have perceived the information's utility to be greater than what a real implementation could reasonably deliver. The ways in which a rendezvous application might use position data are influenced by the nature of the technology used.[7] So, our study provided valuable data given a what-if scenario of consistent, accurate updates, which must be qualified against actual implementations' capabilities.

Marked-Up Maps

The first Marked-Up Maps prototype, which we described earlier, permitted point-and-click interaction with individual points on a map. This was suitable for information access tied to clearly visible landmarks on a map or for providing overview information pertaining to a map grid square. However, we wanted to explore more complex interactions, including filtering information by selecting categories on the map, tracing between two points on a map, and circling map regions. These interactions are suitable for map interaction in general and for expressing queries with a shared map in a group navigation setting. We restricted our evaluation to use by a single user, however, feeling that we could examine group interaction after evaluating the basic techniques.

In this study, the prototype gave visual and audio feedback indicating the selection operations performed but didn't retrieve information. Our software logged whether a particular interaction matched an expected approach for a given task. We gave the user an indication of the items selected, by highlighting map icons, drawing paths between map icons selected in sequence, or highlighting selected map grid squares (see figure 4). We call the prototype a functional approximation not simply because it was

Figure 4. The second Marked-Up Maps prototype: (a) A paper map used with the prototype. (b) One way the prototype provided feedback was to highlight selected map grid squares in red.

incomplete but, more importantly, because the technology limited the granularity with which it captured user interactions. This is similar to a location-based application prototype that uses cell-tower positioning, where the final implementation will have greater positional accuracy.

The granularity of the RFID grid we could create was very coarse, allowing one tag per 30 mm square without interference. This obviously prohibited interactions with finer regions—for example, tracing a walking path as it winds along a map. We considered several alternatives, including using larger or coarser maps, magnetic tracking, and vision-based methods. Ultimately we decided that RFID, despite its obvious accuracy limitations, offered the most direct path for an early prototype evaluation.

Although the granularity was poor, we were able to mock up interactions by detecting certain coarse interactions and interpreting them in software as finer interactions appropriate to a given task. For example, the prototype's software interpreted a sequence of RFID tag reads that corresponded (roughly) to a map's center region as a request for information about a "pedestrianized" area, which was generally contained in that region. The prototype interpreted reading a sequence of tags that contained or skirted a given street's boundary as a swipe along the length of that street if this constituted a suitable interaction for a given task.

The participants underwent brief training to become familiar with the interaction techniques that the prototype supported. They then performed several tasks that didn't specify which interaction technique to use. In general, they didn't remark that selection operations were difficult or unnatural, and as a group they predominantly used the same interaction techniques to perform specific tasks.

This ran counter to the results of an earlier "make believe" evaluation, where we presented the same maps, gave participants handhelds that were turned off, and simply asked them to demonstrate how they might interact with the paper maps to retrieve answers to questions provided to them. During that evaluation, participants envisioned point-and-click as the predominant (and sometimes only) interaction style. The semifunctional prototype created a level of realism and coherence that the make-believe prototype couldn't. So, we could assess receptivity to interaction styles that didn't appear obvious to the participants in the make-believe evaluation.

The functional-approximation approach restricted how we designed our study and analyzed our findings, however. On the map of Halifax we highlighted a grid square corresponding to each RFID tag. The squares were regularly spaced, so they didn't represent specific landmarks (see figure 4). When we asked the participants to trace a path along a set of streets, the visual feedback indicated that they had selected the grid squares that contained the streets in question. If the study's purpose had been to determine optimal visual feedback on the handheld device during selection operations, this prototype implementation would have been suboptimal. Regardless, the feedback is clearly insufficient for tracing routes. When we were devising this study, we needed to decide whether to include tasks such as route tracing for which the prototype might not provide the best solution. The prototype's limitations regarding a subset of the activities might cloud the participants' overall impression. However, we ultimately decided to include such tasks in our evaluation because we were interested in collecting participant evaluations about the interaction technique as an approach in general. We adjusted the order in which participants conducted activities to mitigate any negative impressions given by the path feedback.

The prototype didn't go beyond providing visual and audio feedback because we intended to explore receptivity to interaction techniques, not to evaluate a complete system. If a participant assumed that the desired information would be easily retrieved from a selection, it seems reasonable to expect that he or she would be positive toward the related interaction technique. Needless to say, it's difficult to evaluate something that doesn't actually fulfill the function it's meant to perform. Despite encouraging results in relation to interaction techniques for selection operations, we needed to be careful to not conclude that the system is on the right track in terms of the entire workflow.

Lessons and recommendations

As has been discussed here and elsewhere,[4] Wizard of Oz and other techniques requiring monitoring or following participants can be difficult to manage in mobile environments and can limit realism. For example, the semi-functional and incomplete Marked-Up Maps prototype was better suited to focused evaluation in a stationary location. However, the Wizard of Oz implementation in the Rendezvous study still permitted mobile evaluation.

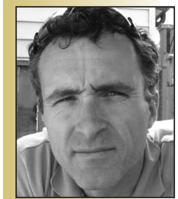
Determining appropriate levels of fidelity and granularity in a prototype is critical before developing and evaluating it. Indeed, rapid prototyping tools for pervasive computing often facilitate this by allowing the introduction of error.[8,9] While it's tempting to use the technology you plan to deploy in your prototypes, the technology can often require further work before being ready (for example, combining RFID with other techniques might improve granularity, and better algorithms might yield better positional accuracy in a location-aware system). On the other hand, as we mentioned before, using a Wizard of Oz approach or an alternate technology might introduce expectations that aren't attainable with the deployed technology. If the capabilities suggested by the prototype are beyond even the far-term capabilities of a real system, the results will have limited practical utility.

Pretending that a prototype is more accurate than it really is poses its own difficulties. As sometimes occurred with Marked-Up Maps, participants can pick up on cues indicating that the system is less accurate than you or the interface suggest. A possible consequence is that the participant simply presumes the system will work "as advertised," which complicates evaluation of user interactions and impressions. Similar difficulties can occur when a Wizard of Oz prototype's mechanics become apparent.

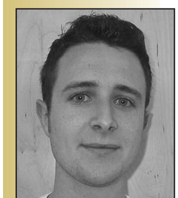
Providing some indication of a system's functionality, even given these limitations, can elicit interactions and impressions that are valuable to designers and researchers. When individuals envision functionality without an interactive prototype, they can be unduly influenced by prior experience and expectations.

We would like to say that experience prototyping and Wizard of Oz prototyping, taken together, constitute a suitable alternative to true contextual evaluation for early prototypes, one gaining realism at the expense of impartiality, the other impartiality at the expense of realism. As our experiences show, however, each pervasive application design poses unique challenges that you must consider when applying these approaches. Regardless, it's often a good trade-off to sacrifice some measure of realism to evaluate early prototypes. Evaluation of early pervasive computing prototypes in context is a pragmatic exercise,

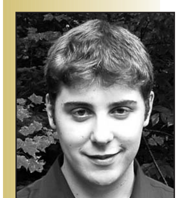
the AUTHORS



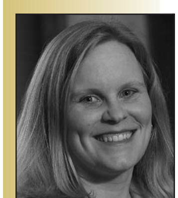
Derek Reilly is a PhD student in computer science at Dalhousie University and a member of the EDGE (Exploring Dynamic Groupware Environments) Lab. His research considers ways to integrate geographic maps with other spatial information in context. He received his BSc in computer science from McGill University. He's a member of the ACM. Contact him at the Faculty of Computer Science, Dalhousie Univ., Halifax, NS B3H 1W5, Canada; reilly@cs.dal.ca.



David Dearman is a graduate student in the Faculty of Computer Science at Dalhousie University and a member of the EDGE (Exploring Dynamic Groupware Environments) Lab. His research focuses on using mobile location-aware devices in social coordination. He received his BCSc in computer science from Dalhousie University. Contact him at the Faculty of Computer Science, Dalhousie Univ., Halifax, NS B3H 1W5, Canada; dearman@cs.dal.ca.



Michael Welsman-Dinelle is an undergraduate student majoring in computer science at Dalhousie University. During the winter term of 2005 he contributed to human-computer interaction research in the EDGE (Exploring Dynamic Groupware Environments) Lab as part of a co-op work term and intends to tackle more research questions in the future for his honors thesis. Contact him at the Faculty of Computer Science, Dalhousie Univ., Halifax, NS B3H 1W5, Canada; michaelw@cs.dal.ca.



Kori Inkpen is an associate professor in the Faculty of Computer Science at Dalhousie University and is the director of the EDGE (Exploring Dynamic Groupware Environments) Lab. Her main research interests involve computer support for collaboration and ubiquitous computing. She received her PhD in computer science from the University of British Columbia. She's a member of the ACM. Contact her at the Faculty of Computer Science, Dalhousie Univ., Halifax, NS B3H 1W5, Canada; inkpen@cs.dal.ca.

but one that is nonetheless informed by general approaches that reflect and adapt to the challenges of pervasive application development. P

ACKNOWLEDGMENTS

Thanks to all researchers involved in the studies discussed in this article: Ritchie Argue, Colin Bate, Vicki Ha, Kirstie Hawkey, Melanie Kellar, Joe MacInnes, Bonnie MacKay, Mike Nunes, Karen Parker, Malcolm Rodgers, and Tara Whalen. We also thank the National Sciences and Engineering Research Council of Canada and Dalhousie University for supporting the research, and the article reviewers for their helpful comments.

REFERENCES

1. G. Abowd and E. Mynatt, “Charting Past, Present and Future Research in Ubiquitous Computing,” *ACM Trans. Computer-Human Interaction*, vol. 7, no. 1, 2000, pp. 29–58.
2. J. Kjeldskov and K. Graham, ”A Review of Mobile HCI Research Methods,” *Proc. 5th Int’l Symp. Human-Computer Interaction with Mobile Devices and Services* (Mobile HCI 2003), LNCS 2795, Springer-Verlag, 2003, pp. 317–335.
3. M. Fleck et al., “From Informing to Remembering: Deploying a Ubiquitous System in an Interactive Science Museum,” *IEEE Pervasive Computing*, vol. 1, no. 2, 2002, pp. 13–21.
4. M. Isomursu et al., “Experience Clip: Method for User Participation and Evaluation of Mobile Concepts,” *Proc. 8th Biennial Conf. Participatory Design*, ACM Press, 2004, pp. 83–92.
5. M. Buchenau and J. Suri, “Experience Prototyping,” *Proc. Conf. Designing Interactive Systems: Processes, Practices, Methods, and Techniques* (DIS 2000), ACM Press, 2000, pp. 424–433.
6. M. Walker et al., “High-Fidelity or Low-Fidelity, Paper or Computer? Choosing Attributes When Testing Web Prototypes,” *Proc. Human Factors and Ergonomics Soc. 46th Ann. Meeting*, Human Factors and Ergonomics Soc., 2002, pp. 661–665.
7. M. Chalmers, “Seamful Design and Ubicomp Infrastructure,” presented at the UbiComp 2003 workshop, At the Crossroads: The Interaction of HCI and Systems Issues in UbiComp, www.equator.ac.uk/var/uploads/M.%20Chalmers%20(2003),%20Seamful%20Design%20and%20Ubicomp%20Infrastructure.2003.pdf.
8. L. Yang et al., “Topiary: A Tool for Prototyping Location-Enhanced Applications,” *Proc. 17th Ann. ACM Symp. User Interface Software and Technology*, ACM Press, 2004, pp. 217–226.
9. R. Hull et al., “Rapid Authoring of Mediascapes,” *Proc. 6th Int’l Conf. Ubiquitous Computing* (UbiComp 04), LNCS 3205, Springer-Verlag, 2004, pp. 125–142.