

A Comparative Prototype Research Methodology

Jonathan Trevor and David Hilbert

FX Palo Alto Laboratory, 3400 Hillview Avenue, Bldg 4, Palo Alto, CA 94304

{trevor,hilbert}@fxpal.com

1. Methodology

Evaluating ubiquitous systems is hard, and has attracted the attention of others in the research community [5]. These investigators, like others in CSCW [3][6], argue there is a basic mismatch between traditional evaluation techniques and the needs posed by ubiquitous systems. Namely, these systems are embedded in a variety of complex real world environments that cannot be easily modeled (as required by theoretical analyses), simulated, measured, or controlled (as required by laboratory experiments). These concerns are shared by Abowd, Mynatt and Rodden, who argue “deeper evaluation results cannot be obtained through controlled studies in traditional, contained, usability laboratory.” [2]. As a result, many investigators have abandoned traditional comparative evaluation techniques and opted instead for techniques adapted from the social sciences, such as anthropology.

We wanted to perform a comparative evaluation similar to a laboratory experiment, but in such a way that we could observe the effects of our design decisions in relatively unconstrained, real world use. This led us to the following process:

1. **Design with alternatives.** Pick a design decision to vary and build prototypes (or configurations of an existing system) to embody multiple design alternatives.

- a. Make the system variants provide the same basic underlying functionality.
- b. Make the underlying system functionality as simple as possible but at the same time compelling and useful enough to produce real usage.

2. **Vary deployment situations.** Deploy the prototypes in varied situations to help answer the question: are observed effects linked to a single situation or are they more general?

3. **Compare and contrast.** Use qualitative or quantitative data collection and analysis techniques to compare and contrast the alternatives and situations.

- a. Try to determine whether observed effects vary depending on the design alternatives, the varied situations, or both.

In addition, the prototypes (or configurations) need to be useful – they need to be used by real people to address real problems and fit in with existing practices. This should be accomplished through techniques such as informal observation of current work practices and problems, or by using more principled studies.

This evaluative method gives us a framework for better understanding a ubiquitous computing system. It goes beyond *designing for use* and promotes *designing for evaluation*. The framework produces a matrix of observations and measurements across situations of use that help us to think about ubiquitous computing design.

2. Case study: Two Personal Interaction Point systems

We set out to design a personal interaction points (PIPs) system for *personalizing shared ubiquitous devices*. We were motivated, in part, by the anonymous interfaces and functionality found in copiers, fax machines, and printers that are so common in hotels business centers, convenience stores, and copy shops. Although there are many personalization features imaginable, we decided that integrating personal computer file access with a shared device’s functions would be a good starting point. The PIPs system embodies this by giving “smart” access to people’s file history *at* shared devices, just like the Windows recent “Documents” menu gives access at desktop computers. The “smart” part of the system was to match file types from the user’s “information cloud” with the function types of devices. For example, if the device is a projector in a conference room then the preferred type is PowerPoint.

2.1. Design alternatives: embedded vs. portable

Many ubiquitous systems, e.g., PalPlates [4], adopt an *embedded* model in which users interact directly with devices embedded in the environment. Other ubiquitous systems, such as CybreGuide[1], adopt a *portable* approach in which users interact with portable devices, such as their cell phones or PDAs. In the PIPs system, we wanted to vary this design decision to learn more about the relative merits of the design alternatives. Thus, we developed two PIPs variants: embedded, which provide an

interface connected to the device itself, e.g., a touch screen; and portable, in which users interact via portable devices, e.g. mobile phones or PDAs.

2.2. Varied situations: presentations, brainstorming, and printing

In order to understand the effects of multiple situations, we took three shared devices that support different tasks in different locations in our workplace, and integrated the PIP system into each, helping users do the same tasks they would normally do in that location, but via a personalized interface that streamlines common activities.



Figure 1. Three places with shared devices in our laboratory: The formal conference room, the informal brainstorming room, and the mailroom.

The podium PC in our formal conference room (figure 1 top) was extended with our “presentation PIP” that provides users with a personal interface for accessing their most recently edited or viewed presentation, which they may then display on the main screen with a single interaction.

Our brainstorming room (figure 1 left) is a much smaller room intended for use by about eight or ten people for informal discussions and brainstorming. We extended an existing plasma display (used for hooking up laptops) with a networked PC running the PIP software, and added a touch screen overlay and wireless keyboard to expedite interactions at the display. Our “brainstorming PIP” provides users with a personal interface for quickly accessing their most recently edited or viewed documents, which they may then view and edit on the plasma display.

Finally, our mailroom (figure 1 right) hosts a multi-function copier device (MFD) with print, copy, scan, and fax functions. The MFD is networked and

centrally located within the building and used by nearly everyone daily. We added a 15-inch touch screen monitor co-located with the MFD, which we connected to a PC running the PIP software hidden in a cabinet beside the MFD. Our “printer PIP” provides users with a personal interface for accessing their most recently edited or viewed documents, which they may then print on the MFD with a single interaction.

2.3. The embedded and portable alternatives

The embedded PIP interface consists of a touch screen on or near the shared device. The PIP is activated when the user approaches the shared device and swipes their ID card over the card reader.

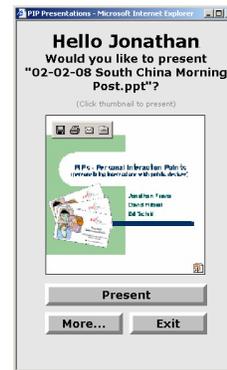


Figure 2. The embedded PIPs “Best Pick” (top) and “Full” (bottom) interfaces.

A PIP Web application then generates the personalized interface by fetching and resolving the shortcuts stored in the user’s recent file list on their PC. The PIP presents a “best pick” interface with the recent file (or files) the user is most likely to want to use at the PIP-enhanced device (figure 2 top). The user may then perform a default action (such as “present” or “print”), by pressing on the document’s thumbnail using the touch-screen provided by the embedded PIP. Files are accessed over the network from their original locations, so users needn’t plan ahead or copy files anywhere.

If the best pick interface does not contain the user's desired document, the user can press the "More..." button to bring up the "full" interface (figure 2 bottom). This allows the user to access virtually any document (via the device) that they have ever accessed on their office PC.

Our comparative prototypes methodology requires us to provide near identical functionality for both design variants, so it was important to design the portable interface to be as similar to the embedded interface as possible.

To use the portable PIP interface, user's point their device's Web browser to the PIP home page. Selecting a PIP-enhanced device takes the user to the portable PIP interface for that device. Standard browser authentication is used to prompt for their NT username and password. Once authenticated, the PIP returns a splash screen and proceeds to fetch and resolve the user's recent file list in the same way it does for the embedded interface. However, using the PIP with a smaller portable device, such as a Pocket PC, required us to make a few cosmetic changes due to the smaller screen dimensions. For example, the file details view (the right-most frame in the full display in figure 2) was separated into its own page.

2.4. Deployment and Observations

The PIPs prototypes were deployed and adopted over the course of several months. The three embedded PIPs were released a year ago, and the portable interfaces released a couple of months later. Ideally we would have deployed the prototype variants simultaneously. However, we believe this didn't significantly affect our findings since many of our users only started using the system after both variants were available.

The initial months were spent debugging the prototypes and increasing the visibility of the PIP-enhanced devices within the laboratory. Gradually, our user base grew as lab members observed the utility of the system as demonstrated by early adopter usage, primarily in our formal conference room. The trend has been toward increased users *and* usage. Approximately 80% of the research staff is now using PIPs, and so far no one who has used it has subsequently stopped using it. The presentation PIP is used for over half of the presentations given in our formal conference room; the brainstorming PIP is used for nearly all documents accessed in our brainstorming room. Finally, the printer PIP is used for less than one percent of all print jobs.

During the past year we gathered feedback from early adopters, late adopters, and non-adopters, and recorded incidents in which PIPs failed to operate as expected by our users or us. These eventually became known as "PIPcidents", and provided much of the fodder for the issues we found.

2.5. Summary of Findings

Our comparative prototyping methodology illuminated five key issues in the design of personalized ubiquitous systems: usability, utility, availability, trust, and privacy. Table 1 summarizes differences between the alternative interface approaches across three situations.

Issues of **usability** and **availability** varied more based on design decision (embedded versus portable) and less by situation. Embedded personal interfaces were more *usable* than portable versions irrespective of the situation, due primarily to their larger display size. Even when run on a laptop, the portable interface still produced more usability problems than its embedded counterpart due to the separation of the user interface from the underlying device. Embedded interfaces were also more *available* than portable interfaces across situations, being less prone to problems of wireless networks, battery life, and having to carry a device with you at all times.

The issues of **utility** and **privacy**, on the other hand, varied not only based on design decision, but also by situation. The portable presentation PIP enabled remote control functionality that was particularly useful in our formal conference room, but less important in the other two situations. Portable interfaces were also better than their embedded counterparts for supporting user privacy. Again, this difference was more marked in our formal conference room, but less marked in our informal brainstorming room and in our mailroom where the likelihood of being overseen was reduced, in part due to the quick nature of activating print jobs. Indeed, it seems that short, less intricate tasks, like printing, were generally well suited to portable devices, whereas embedded, custom, interfaces are much better at supporting more complex tasks.

User **trust** did not appear to vary significantly based on design decision or situation in our experiment. However feedback from visitors indicates that people may feel more comfortable accessing personal data via personal portable interfaces, than interfaces embedded in shared devices, particularly at public places such as Kinko's or a 7-Eleven.

PIP	Situation	Decision	Issues in personalization shared ubiquitous devices				
			Usability	Utility	Availability	Trust	Privacy
Presenting PIP	Formal meeting room	Embedded	++	+	++	+	-
		Portable	+	++	+	+	++
Editing PIP	Small meeting room	Embedded	++	+	++	+	+
		Portable	+	+	+	+	++
Printing PIP	Public mail room	Embedded	++	+	++	+	+
		Portable	+	++	+	+	+

Table 1. Summary of issues for personalizing shared ubiquitous devices

3. Methodology Issues

During the development and use of the comparative methodology we became aware of several issues with its design and application:

- **Technical effort.** The approach calls for alternative designs which require extra time and effort to build. Considerable effort is expended with no guarantee of significant evaluation results.
- **Social effort.** Politically it is often necessary to “champion” a technology to get it used. With our approach, one must champion several similar technologies and not show favoritism in promoting a given alternative over the other(s).
- **Design decisions not atomic.** When you vary a design decision, you are likely varying a number of factors. Determining which particular aspect of the decision caused a particular result can require additional effort. In our example, were the embedded versions easier to use because they were fixed and collocated with the device? Or primarily due to their larger size? Perhaps a combination of both factors? Further evaluation would be required.
- **Varied situations not atomic.** As with design decisions, when varying the situation, you are likely varying a number of factors. Determining which particular aspect of a situation caused a particular result may require additional effort. In our example, were users less concerned about privacy when printing because the interaction time was short? Or because people were typically alone when printing? Or a combination of both factors? Again, further evaluation would be required.
- **Complicated analysis.** Due to the two previous issues, the methodology doesn’t produce simple “A is better than B” answers. This is due to the difficulties in identifying which particular aspects of design decisions and deployment situations caused

particular observed advantages or effects. Thus we measured a number of different attributes in our evaluation to help tease out some of the factors causing observed differences. We identified several broad issues which seemed different between the systems and used those to focus an examination on why they differed. Follow-up evaluation (more focused user interviews or detailed experiments) might be performed afterwards to learn more about particular effects of interest.

- **Generalization.** Related to the above points, we clearly need to work on techniques for more clearly characterizing the design decisions and deployment situations in order to generalize our findings. Is real-life design decision A in situation X similar to experimental design decision B in situation Y?

While we are only beginning to explore this methodology, we feel it holds great promise for performing HCI-style comparative evaluations on UbiComp and CSCW systems.

References

1. G. D. Abowd, C. G. Atkeson, J. Hong, S. Long, R. Kooper, M. Pinkerton. Cyberguide: a mobile context-aware tour guide, *ACM Wireless Networks*, 3, pp. 421-433, 1997.
2. G. D. Abowd, E. D. Mynatt and T. Rodden. The Human Experience. *IEEE Pervasive Computing*, March, pp.48-57, 2002.
3. J. Grudin. Why CSCW Applications Fail: Problems in the Design and Evaluation of Organisational Interfaces. *Proceedings of CSCW '88*, 1988.
4. J. Mankoff and B. Schilit. Supporting Knowledge Workers Beyond the Desktop with Palplates. In *Proceedings of CHI 97*, 1997, pp. 550-551.
5. J. Scholtz and A. Smailagic. Workshop on Evaluation Methodologies for Ubiquitous Computing, UbiComp 2001. Published in *SIGCHI Bulletin*, Jan/Feb 2002.
6. M. Twidale, D. Randall and R. Bentley. Situated evaluation for cooperative systems. *Proceedings of CSCW 94*, 1994.