



Extremely Rapid Usability Testing

Mark Pawson

Business Analyst
S.i.Systems
Calgary, Alberta
Canada
markpawson@xplornet.com

Saul Greenberg

Computer Science
University of Calgary
Calgary, Alberta
Canada
saul.greenberg@ucalgary.ca

Abstract

The trade show booth on the exhibit floor of a conference is traditionally used for company representatives to sell their products and services. However, the trade booth environment also creates an opportunity, for it can give the development team easy access to many varied participants for usability testing. The question is can we adapt usability testing methods to work in such an environment? Extremely rapid usability testing (ERUT) does just this, where we deploy a combination of questionnaires, interviews, storyboarding, co-discovery, and usability testing in a trade show booth environment. We illustrate ERUT in actual use during a busy photographic trade show. It proved effective for actively gathering real-world user feedback in a rapid paced environment where time is of the essence.

Keywords

Discount usability methods, rapid usability testing, trade shows

Introduction

Traditional usability testing typically occurs in a laboratory-like setting. Participants are brought into the test environment, a tester provides tasks to the participants, and the participants are instructed to “think aloud” by verbalizing their thoughts as they perform the tasks (e.g., Dumas & Reddish, 1999; Nielsen, 1996). Observers watch how the participants interact with the product under test, noting both problems and successes. While a typical usability test normally takes at least one hour to run through several key tasks, it can take many days or weeks to set up (e.g., lab and equipment set up, protocol preparation, recruitment, scheduling, dealing with no-shows, etc.). The key problem is that it may be quite difficult and/or expensive to motivate people—particularly domain experts—to participate in such a study. While this can be mitigated by running the test in the domain expert’s workplace, this introduces other significant problems, such as disruptions to the expert’s actual work.

Another possibility is to use a trade show as a place for conducting usability tests, especially for new versions of a product that would naturally fit a trade show theme. We can consider the benefits of a trade show in light of Dumas and Reddish’s (1999) following five characteristics of usability testing:

1. The primary goal to improve the usability of a product...
2. Participants represent real users,
3. Participants do real tasks,
4. You observe and record what participants do and say and
5. You analyze the data and recommend changes

A trade show emphasizes characteristics 1, 2, and 3. Characteristic 2 is the one that is maximized: there is a plethora of potential participants, all very real users with domain expertise, not only present but likely willing to participate in the usability test. They should be highly motivated to try out, and thus test, new product versions. Their attendance means they have a large block of time for doing so. Next, a trade show setting sets the scene for characteristic 1 because trade shows largely concern advertising, familiarizing, and ultimately selling a product to potential customers. Product features, usefulness, and usability dominate discussions between participants and those manning the booth. For characteristic 3, participants are engaged by the theme of a trade show, they could easily reflect upon the actual tasks that they would want to perform on a product or critique the tasks they are being asked to do. In turn, the feedback gained is likely highly relevant to real-world use.

Yet there are issues. A trade show is not a laboratory, nor is it a workplace. Trade shows are crowded and bustling venues, where vendors compete with others to attract people to their booths. A trade show exhibit booth is a hectic, noisy, cramped space that exists for three days and could be visited by 500 people or more. Booth visitors can be users, competitors, students, or future customers. Each visitor may spend anywhere from one minute to 60 minutes in a booth. Distractions are rampant. This is not a typical usability test environment! This makes characteristics 4 (observe and record) and 5 (analyze) more problematic for the evaluator and constrains the kinds and number of tasks (characteristic 3) that can be done. Yet for companies with limited time and resources to get their product to market, a trade show could offer a realistic way to gather a broad brush of domain experts in one place for product testing.

Of course, there are evaluations methods within human computer interaction (HCI) that others developed for time and resource limited environments (e.g., Bauersfeld & Halgren, 1996; Gould, 1996; Marty & Twidale, 2005; Millen, 2000; Thomas, 1996), but none specifically address the trade show setting. Gould (1996) was perhaps the earliest advocate of rapid testing. He describes a plethora of highly pragmatic methods that let interface designers quickly gather feedback in various circumstances. His examples include placing interface mockups in an organization’s hallway as a means to gather comments from those passing by and continually demonstrating working pieces of the system to anyone who will take the time to watch. The advent of quick and dirty usability testing methods in the mid 90s formalized many of these processes. Each method was an attempt to decrease the cost of the test (time, dollars, resources, etc.) while maximizing the benefit gained (e.g., identifying large problems and effects, critical events, and interface compliance to usability guidelines, etc.) (Nielsen, 1994;

Thomas 1996). Other methods were developed to specific contexts. For example, Marty and Twidale (2005) described a high-speed (30 minute) user testing method for teaching, where the audience can “understand the value of user testing quickly, yet without sacrificing the inherent realism of user testing by relying solely on simulations.” Millen (2000) discussed *rapid ethnography*, a collection of field methods tailored to gain a limited understanding of users and their activities within the constraints of limited time pressures in the field.

No method specifically addressed running rapid usability tests in a busy trade show or conference exhibit hall booth. The question remained, how can we use the trade show as a place for conducting usability tests? Consequently, our goal was to see if we could adapt and modify existing usability testing methods to the trade show context, which we called *extremely rapid usability testing* (ERUT). Our experiences with ERUT involved a pragmatic combination of HCI evaluation techniques: questionnaires, co-discovery, storyboarding, and observational think-aloud tests. It was an example of taking formative testing methods and applying it to a particular context of use. We wanted to exploit the “best” of each method, i.e., the portion that delivers the maximum amount of information within the severe limitations of the trade show. ERUT is not a formal or exhaustive usability evaluation of a product, nor a replacement for other methods. Rather, ERUT applies and mixes various informal discount methods to provide insights into the usefulness and usability of primary product features.

ERUT developed opportunistically. This paper’s author, Mark Pawson, and another colleague were invited by Athentech Inc. of Calgary Alberta to attend the PDN PhotoPlus show in New York to perform rapid usability tests on the Perfectly Clear[®] digital imaging enhancement software. Pawson already worked as a usability evaluator, and both he and his colleague were experienced in working trade booths from a marketing perspective. We developed ERUT to quickly gather real-world feedback about the usability and usefulness of this product and to shed significant light on whether Athentech’s *unique selling proposition* resonated with the customer.

In the remainder of this paper, we describe our experiences developing and using ERUT to evaluate Perfectly Clear[®] at the PDN PhotoPlus trade show. We caution that ERUT as described here is a case study of our experiences and the lessons we learnt, rather than a rigid prescription of how to do usability testing in a trade show environment. That is, it can be seen as a starting point for practitioners to adapt usability testing to their own trade show settings.

The Product and Context

Athentech states that Perfectly Clear Pro[®] is a digital image enhancement software designed to correct a digital photo to match what the human eye sees when the picture was taken. Without getting into technical details, Athentech developed a process that overcomes camera limitations and produces photos that yield what the photographer saw when capturing the image.

Athentech licenses this technology to photographic labs and to industry leaders such as Fuji, Blacks, Ritz, and Walgreens for use in kiosks and mini-labs. They also wanted to enter the professional consumer market. To this end, Athentech regularly attended trade shows to understand the problems photographers face with digital imaging and with existing software tools on the market. They then developed Perfectly Clear Pro[®] as their first venture into developing a product for the professional and serious amateur photographers.

In our specific case, Athentech was keen to take an alpha version of Perfectly Clear Pro[®] to PhotoPlus, a major trade show and exposition whose tag line is “...to be on the cutting edge of what’s happening in photography and imaging” (<http://www.photoplusexpo.com>). However, Athentech had not yet performed any usability evaluations. They believed the show represented a tremendous opportunity not only to get their product in front of many potential customers in a very short time but to try to understand where the alpha version succeeded and failed.

From prior experiences, we knew that running usability tests in a booth would be quite different from the usual evaluation setting.

- The trade show had strict daily closing times, which meant testing after show hours would not be possible.
- The environment was noisy. While Athentech had chosen a closed booth with a section cordoned off for testing, cordoning was done via curtains.
- Participant selection would be haphazard, as it depended on who we could attract from the general conference milieu.
- Testing time was very limited. Past experience with booth visitors indicated that having 15-20 minutes of a participant's time would be generous. Although some participants would perhaps stay longer, most would fit this in-between talks and visits to other booths.
- Time to immediately reflect on particular study results was limited due to the need to process as many people as possible within the three day duration of the show.
- From the participant's perspective, usability testing was only one purpose—the lesser purpose—of the booth. When a visitor stressed business needs over a desire to be a test participant, the tester would have to rapidly switch from wearing a usability testing hat to a sales hat.

A testing regime has to be fluid in order to respond to these constraints. Consequently we designed ERUT to focus on the following two primary objectives:

- Assess the usefulness of the core functionality of a product, i.e., was the product's unique selling proposition solving a problem that a majority of customers wanted solved?
- Find major usability problems in the core functionality.

While this meant that some aspects of the software would be ignored, we hoped that ERUT could determine the usefulness and usability of the core product.

Methodology Details

The following sections discuss the booth setup; recruiting participants; questionnaires; choice of tasks; co-discovery, think aloud, and active intervention techniques; and storyboarding for recording results.

Booth Setup

The trade booth doubles as both a marketing venue and the usability testing area. While it is possible to have two separate booths, we believe a single one is best as it is the product marketing that attracts the participants (discussed shortly). Still, it is important to isolate the testing area from the direct flow of the convention crowd, perhaps by partitioning the booth into two areas: an outer booth for marketing and an inner booth for testing. Without an isolated quieter area, the evaluator runs considerable risk of introducing interactions and distractions in the booth between test participants and those wandering in and out of the booth (IXDA 2007).

In our case, PhotoPlus attracted huge crowds with over 27,000 registered participants. To adjust the flow of potential participants and to isolate the test area, we set the booth walls up around the outside perimeter of the allotted booth area assigned to PhotoPlus. The outside of the booth walls were hung with promotional posters and sample pictures of Perfectly Clear[®] technology, as illustrated in Figure 1. We then created a doorway into the inner booth, which became the test area as illustrated in Figure 2. As discussed below, the Athentech marketing representative would then feed participants through this doorway when we were able to receive them.



Figure 1. The booth's exterior, used for product promotion and marketing. Note the doorway to the interior testing area on the right.



Figure 2. The booth's interior, used as a testing area.

Recruiting Participants

The trade show offered ease of access to a large variety of domain experts and potential customers in one place. The question was how do we recruit these people given the large number of other booths competing for their attention?

In our case, the attractant was the pictures that hung on the booth wall exteriors that displayed the before and after effects of the Perfectly Clear[®] technology (Figure 1), and the unique selling proposition delivered by the Athentech representative working the front of the booth. The Athentech representative served as our gatekeeper. He invited interested potential customers to test the product, while controlling the flow into the testing area.

Interested attendees typically asked a booth representative for a demonstration. While many booths provided such demonstrations, our representative explained that the product was still in its early stages and that only those people willing to participate in usability test could try it. Those who volunteered to participate in usability testing were then invited into the booth on a first come, first served basis. Participants felt that they were in control of this process, for it came out of their desire to try the system. To make this work, much of the preliminary process that precedes a usability study was discarded. For example, we did not use written consent forms, nor did we offer incentives to have people participate in the usability testing (although we did give participants gifts of all-natural chocolate from the Amazon rainforest). Certainly, the issue of consent has to be revisited both to inform the participant more clearly and for organizational liability; the question is how to do such consent effectively within this context.

Of course, we could not handle all possible participants due to time constraints. Yet those who could not participate were not necessarily lost opportunities. We scanned in contact information from the badges of several hundred trade show attendees who were interested in trialing (and thus evaluating) a beta copy of the product at a future time.

Questionnaires

We originally planned on a short pre-questionnaire and an optional post questionnaire (e.g., a satisfaction or a desirability survey). We knew that time would be short in the booth and that participants would be eager to get to the product, so we wanted the questionnaire to be equally short. Thus we focused only on a few key questions that the company considered critical.

Athentech's previous research had already validated that Perfectly Clear[®] was aligned with customer goals. Their concern was with the offerings of recent competitive products on the market. Athentech felt that those products offered a different workflow and unnecessary functionality. Athentech also thought other vendors had understated the limitations of the digital camera in capturing true images. Given this, we targeted our pre-questionnaire to simple demographics (if participants were professional or serious amateur photographers), what software tools they were currently using for their work, and what they were using these tools for.

However, there were tradeoffs. Athentech also wanted to collect additional user feedback on various topics that would help guide their future software development. This would have dramatically increased the size of the questionnaire. We were concerned that customers would be turned off; they were drawn to the usability test (which was in the spirit of trade show demonstrations) but not to the barrage of questions both before and after the test. We found it challenging to balance the questionnaire so that it met both business and testing needs while respecting the customers' short timelines and interests. As discussed later in our "Lessons Learnt" section, flexibility was the best approach. Instead of requesting this extra information as part of the written questionnaire, we worked the questions into our conversation with participants while they were doing the task. We were opportunistic: we asked questions when they fit into the flow of activity, but in the interest of time not all questions were asked.

We also found that our post-test survey questionnaire did not work in the context of the booth. The questionnaire did not fit with the natural rapport of a trade show booth. As one participant said "everything you have done up to now has been great, but this just turns me off."

Choice of Tasks

We develop three tasks ahead of time that were both unique and representative of problems we believed that potential customers wanted solved and that incorporated the unique selling proposition of Perfectly Clear[®]. This was a modification of an idea used by Chauncey Wilson for testing in a trade show booth (personal communication, 2007). We had planned to let participants select the most personally interesting one of these three tasks to do. We thought the choices made would give us insight on what parts of the product the participant perceived as the most useful.

However, we decided that this approach was not the best one. First, the alpha release of Perfectly Clear[®] was not robust enough to allow people to actually do some of these independent tasks. More importantly, Perfectly Clear[®] was targeting a specific task workflow, cull and image correction of photos. Athentech was in part positioning itself against its competitors who (Athentech believed) had lost sight of this basic customer need by adding layers of complexity and functionality. Consequently, we decided to concentrate only on a core task that addressed this specific workflow. If that could not be done by people to their satisfaction, then it wouldn't really matter how well they could do other tasks with the system. Therefore we spent time with Athentech learning about the specific problems photographers faced with image enhancement and how this was addressed by Perfectly Clear's[®] workflow. From this we created four interrelated scenarios in a photographer's language that we felt were both representational and motivational. These tasks were originally written down on 4 x 6 cards and were to be given to the participants as they completed each task. However, as in the questionnaire, we found the best way to introduce the task was as part of an informal conversation rather than by script. Hence the exact language used to introduce each of the four tasks varied between participants.

While the above may sound like normal task selection and debugging, we want to stress that the short time line forced us to reconsider our tasks. We would likely have time for people to do only a single task, and we needed to ensure that the results were extremely practical.

Co-discovery, Think Aloud, and Active Intervention

We were concerned that the trade booth could create an intimidating atmosphere for usability testing. We did not know ahead of time how the booth layout would affect participant privacy and distraction, which in turn would hamper the concentration of a single participant asked to “think aloud” while completing a task. We decided to use co-discovery, where two participants work together to complete a task. Co-discovery yields higher quality verbal communication between paired participants than single participants. The pair typically converse for their own benefit to complete the task, as opposed to a single participant who is communicating solely for the test facilitator’s benefit.

In the trade show context, we felt it unreasonable to pair strangers. Instead, we looked for people who visited the booth with a friend or associates and encouraged them to be our participants. Still, we did use single participants if no pair was around at the moment. In these instances, and given the predicted short test cycles, we used active intervention in order to elicit high quality think aloud comments. Active intervention was also advocated in a Web discussion forum on usability testing at conferences (IXDA, 2007). We were somewhat surprised at how well this worked. Only once did we have to ask a participant what they were thinking, all others proved textbook examples of the think aloud technique. We surmise that this is the result of the informality of our private testing area, the relaxed trade show atmosphere of the attendees, and participants’ keen interest in the product.

In practice, we gleaned equally high quality think-aloud and co-discovery comments from both individual and paired participants. We certainly observed the engagement of paired participants with each other as the research has reported. However, we also found that it was quite common for one participant to break off his conversation and attention to the task. The participant would explain her thoughts to us or ask a question, while the partner carried on alone. We used active intervention on both the single and paired participants to work in guiding questions at appropriate times.

Storyboards for Recording Results

Recording test results in the fast pace, noisy atmosphere of the trade show raises other challenges. We used a modification of an HCI discovery technique described by McQuaid, Goel, and McManus (2003) to shadow and record the “story” of library visitors. They took pictures of the visitors as they pursued their activities. They printed these pictures and overlaid acetate sheets to record their notes of what they observed. Then, they compiled these into storyboards that they hung on a wall and displayed to stakeholders.

In a similar way, we used hardcopy screen shots of Perfectly Clear[®] to record the story of the paths the participants took in exploring the task. To clarify, storyboarding is a prototyping technique usually used to describe an interface sequence to others. Instead, we used storyboarding for note-taking, where the visuals and annotations described the primary actions a person actually did. We did not use videotapes or screen-capture software for recording the usability results, as we would not have had the time to revisit, analyze, and reflect on these recordings. As well, we were looking for high-level vs. detailed effects. It was unclear if video analysis was worth the effort. The advantages of paper storyboards are the ease in taking notes by simply circling or numbering areas visited, adding annotations as needed, and—perhaps most importantly—the immediacy of the result. The storyboards helped us collate our notes at the end of the day and perform our analysis without having to wade through hours of video tape. However, the storyboards are by no means neat, as annotations were made in a rushed pace. Notes on interesting observations, comments made by the participants, answers to questionnaires could all end up on a storyboard and these could be hard to decipher days later. Ours had to be looked at on the same day while our memories were still fresh in our mind. Also, unlike McQuaid, Goel, and McManus’ storyboards, ours were far too messy to show to stakeholders.

Lessons Learnt

While every trade show and usability testing needs differ, we offer the following lessons learned for others to consider within their context.

Easy access to domain experts and potential customers. Perhaps the biggest advantage of ERUT over a standard usability testing methodology is the ease of access to a large variety of domain experts and potential customers in one place. There is no time spent recruiting participants, dealing with the logistics of scheduling, or losing time due to no-show participants. These issues simply do not exist. A trade booth, if designed well, is a natural attractant for people. People are at a trade show because they want to be, and they come into a booth because they are interested in the product. Recruiting these people as study participants is just a matter of suggestion.

Business comes first. In a trade show environment, the business need comes first. Most companies enter trade shows for marketing, not for testing. More importantly, trade show participants are there to see products, not to test them. Thus one should not expect to do rigorous usability testing in such an environment; incomplete questionnaires and tasks are the norm, and participants may shift their attention to their personal needs vs. keeping strictly to the test regime. Yet this shift of attention is also an opportunity, as it creates a type of contextual interview around the topic of user and business needs while running the test task (it is contextual in the sense that the trade show offerings are often part of the conversation). In fact, our experience from this was that a trade show booth might be the next best thing to observing users in the context of their real environment because they are there for themselves, seeking real solutions to problems they have, and they are primarily in the booth for their own personal gain. The result is very rich customer input on their needs.

Casual conversation over scripts and questionnaires. The best way to engage participants was to drop the usability script and questionnaires; we used casual conversation instead. In our case, participants had a real need for automatic batch correction of their photos. Event photographers in particular were in the booth because they wanted to know how Perfectly Clear[®] would save them time doing hundreds of image corrections and allow them to get back to their jobs—shooting photos. They were captivated by the message that they had heard from the Athentech representative and were keen to see the software. Introducing ourselves by giving the standard “thank you for participating in our usability test...” patter and then presenting them with consent forms and a pre-test questionnaire was cold and robotic and did not fit the pace of action. Instead we worked both the business questions and the task into an exploratory conversation. This immediately engaged them, showed respect for their time, and worked with the natural flow of a trade show environment. Participants wanted to talk shop, not be treated as a test subject. They were there to get answers, not to be asked questions. By being very familiar with the questions we wanted to ask, we looked for opportunities to introduce them as part of a conversation during the testing. This was probably the greatest value of the questionnaires—they became our talking points. The questionnaires helped us pick up on important points made by the participant that otherwise could have gone unnoticed unless one is a domain expert in photography. Of course, this comes at a cost, the loss of a script means that the process is not as repeatable. Different words (and different evaluators) may motivate people differently and large chunks of the script may be omitted. This also implies that the collected data is better seen as samples rather than a consistent outcome based on repeatable instructions and tasks.

Tasks need to be meaningful. The actual tasks done by participants and how they are introduced may also deviate from the script. The trade show setting meant that we needed to introduce the task in a way that was meaningful to the participant. In one case, we had a pair of participants who were looking in detail at a photograph and expressed a desire to make the red colors “pop out.” Perfectly Clear[®] corrects photos back to the true colors; artistically enhancing colors (typically done using other products on the market) is not a feature. However, the software does offer an export function. Thus we changed our task on the fly to fit the participant’s expectations and workflow. Originally, our final task read, “Now you have completed your enhancements, pick your three best photos and store them as high quality JPEGs in a folder of your choice on your computer.” We turned their comment around and

simply asked them, "How would you get that photo into the software of your choice to pop out that red?"

The test requires a narrow focus on core issues. Focusing on core issues is critical, not only because time is short (Bauersfeld & Halgren, 1996; Millen, 2000; Thomas, 1996), but because it is likely those core issues will engage participants. Another advantage of the narrow focus is that it requires all stakeholders to define what the core functionality of the product is and what they hope to gain from usability testing in such an environment.

Interruptions are the norm. Even though participants were in a screened inner booth, interruptions happened and had to be accommodated. An example includes participants answering their cell phone. As well, some participants had to leave partway through the test due to conference talks or catching the last train home. Unlike a normal usability test, we could not expect people to set aside a fixed block of time solely for our purposes.

Participants perceive the test primarily as a demonstration. The trade show is a place to gather materials and see demonstrations. Even though we told people they were in a usability test, they still thought of it as an opportunity to try out the system, i.e., they did not really dwell on the fact that they were in a usability test. In one case, a participant responded to a cell phone call from a colleague by saying "Yeah, I'm in a demo right now. I want to buy this software, ok bye." To keep in this spirit, our final question was "would you buy this software?" As well, participants had the opportunity to sign up to get beta-releases of the system.

Tag teaming and active intervention. We found the best sessions were when the two experimenters were able to tag team each other rather than working alone. Although we tried working alone, there were times when note-taking disrupted the natural conversation with the participant. Key observations could have been missed, and the participant (whose time is precious) had to wait for the note-taker to catch up. Tag teaming allowed us to engage and disengage with the participant. One of us would write notes while another would pick up with a thread of interest. Tag teaming was a better fit to the trade show atmosphere, where we could engage participants in friendly conversation rather than sitting back quietly and watching. This active intervention by a team meant that participants were always being observed, that notes were being taken, and that they could talk to us any time.

Test time is variable. We originally felt that 20 minutes was the maximum time that we could expect from any participant. In practice, and somewhat surprisingly, most participants stayed much longer than that because they became engaged with the system. We allowed people to stay longer than planned when this happened. This also meant that strict scheduling could not be done. Instead, our "gatekeeper" would feed us participants as we were able to receive them.

Participant flow must be regulated. Because no scheduling is done, we needed some way to control the flow of participants into the test area. In practice, there were times in the booth where participants were let in too soon after a test has been completed, leaving us scrambling to get prepared (we needed about ten minutes between each test to collate our results, finish up any notes, and get the material ready for the next test). The problem was that the gatekeeper was busy with his own needs (marketing) and sometimes used the departure of a participant as an (incorrect) cue that we were ready for the next one. It would have been helpful to have had a green and a red flag by the doorway for the gatekeeper's benefit (red meant busy, green is ready for more test participants).

Conclusion

ERUT is a valuable adaption and combination existing methodologies to use in public trade show situations where a company exhibits its products. A wide array of actual and potential customers are coming to these exhibits of their own accord. Being able to get a product in front of them for their evaluation is very attractive, and for some companies may be the only chance to run usability tests with true domain experts. ERUT can be both effective and inexpensive. It can provide guidance to what product features really matter to customers and where major usability (and usefulness) problems exist. This information can inform business aspects of the software (i.e., the validity of the selling proposition), software development (i.e., features to include, exclude, refine), and—most importantly as usability practitioners—those key areas of the product that should be evaluated using more formal HCI techniques. ERUT can also validate

learning gained from rapid field methods such as contextual interviews, or from other methods such as heuristic reviews (Thomas, 1996), or even the external validity of laboratory-based usability test results.

When working a trade booth the participant is in control of the time and its use. Expect interruptions and be fluid enough to change from a usability tester's hat to a business hat. Remember that participants are in the booth for their benefit first, so your rapport with them in regards to questionnaires and test tasks must engage them on their level. When this is done well our experience is that extremely rapid usability testing can be an effective way of gathering user feedback in a trade show environment. As a rapid method to get in front of customers and elicit feedback on product direction, it is excellent.

There are cautions. As Thomas (1996, p.112) notes, results from quick and dirty methods are "illustrative rather than definitive." This method can provide insights only into usability issues. The results are not gospel and thus one must guard against the project stakeholders who treat this as the only evaluation procedure (especially if the results are very positive). Similarly a trade booth environment can generate its own excitement and could give a false sense of product success. There are also valid arguments against discount usability methods (Cockton & Woolrych, 2002). Certainly, we need more experiences and debate within HCI regarding collecting user feedback in such environments.

Practitioner's Take Away

The following are advantages and disadvantages of performing extremely rapid usability testing (ERUT) at trade shows.

Advantages

- The testing provides for a light weight, rapid gathering of good quality user feedback without a lot of overhead for preparation and running of tests.
- There is a narrow focus on business goals and core functionality that produces valuable insights.
- There is easy access to a broad brush of credible users.
 - Access to domain experts is easy.
 - There are no "no show" participants.
 - The data can be easily collected in a user database for future tests.
 - Participants are in the booth for their benefit first, which yields rich customer input. This could also be a disadvantage if it generates false excitement.
- The method is very fluid. Company representatives must adapt to change to suit the situation.

Disadvantages

- The focus tends to be narrow.
 - This type of on-the-fly usability testing does not look at all of the product's capabilities. Because of the time constraints, only a few aspects of the product can be evaluated.
 - Core tasks tested in isolation may not represent what happens if that task were performed in the context of a complete application workflow.
- The trade show environment is rapid and hectic.
 - Key observations can be lost because of interruptions.
 - Questionnaires and storyboards can be reduced to scribbles because of the time constraints and the desire to quickly capture as much data as possible.
- Participants are not in their natural environment where they would use the product.
 - Observations are not made in context of real work.
 - Participants are in a trade show frame of mind. They could be affected by the excitement in the booth.

Acknowledgements

Mark Pawson thanks his colleague and friend Marc Shandro in challenging him to join him at PhotoPlus. Most of all, thanks to Athentech Technology Inc.'s president, and Mr. Pawson's former mentor, Jim Malcolm and vice president Brad Malcolm for giving him the opportunity to be engaged in such a unique project.

References

- Bauersfeld, K. & Halgren, S. (1996). "You've got three days!" Case Studies in Field Techniques for the Time-Challenged. In D. Wixon & J. Ramey (Eds.) *Field Methods Casebook for Software Design* (pp.177-195). New York, NY, USA: John Wiley & Sons.
- Cockton, G. & Woolrych, A. (2002, Sept/Oct). Sale Must End: Should Discount Methods be Cleared off HCI's Shelves? *Interactions*, 13 -18.
- Dumas, J.S. & Redish, J.C. (1999). *A Practical Guide to Usability Testing*. Great Britain: Cromwell Press.
- Gould, J.D. (1996). How to design usable systems. In R. Baecker, J. Grudin, W. Buxton and S. Greenberg (eds.) *Readings in Human Computer Interaction: Towards the Year 2000* (pp. 93-121) San Francisco, CA: Morgan-Kaufmann.
- Marty, P.F. & Twidale, M.B. (2005, July). Usability@90mph: Presenting and Evaluating a New, High-Speed Method for Demonstrating User Testing in Front of an Audience. *First Monday* 10(7).
- McQuaid, H.L., Goel, A., & McManus, M. (2003). When You Can't Talk to Customers: using Storyboards and Narratives to Elicit Empathy for Users. In *DPPI 2003* (pp. 120 -125) ACM Press.
- Millen, D.R. (2000). Rapid Ethnography: Time Deepening Strategies for HCI Field Research. In *DIS 2000* (pp. 280-286) ACM Press.
- Nielsen, J. (1994). *Usability Engineering*. San Francisco, CA: Morgan Kaufmann.
- Thomas, B. (1996). 'Quick and dirty' usability tests. In Jordan, P.W., Thomas, T., Weerdmeester, B.A., & McClelland I.L., (Eds.) *Usability Evaluation in Industry* (pp. 107-114). London: Taylor & Francis Ltd.
- IXDA forum thread, various authors. (2007, July 31-September 7) Usability testing at conferences, <http://gamma.ixda.org/discuss.php?post=18865>.

About the Authors



Mark Pawson
Mark Pawson has spent 25 years in Oil/Gas and Geomatics software development service sectors, with the past 15 years leading software support, testing, and user interface design teams at IHS Inc. He currently provides expertise in QA management and process and design input at S.i.Systems in Calgary.



Saul Greenberg
Saul Greenberg is a full professor of computer science at the University of Calgary. He holds the NSERC/iCore/SMART Technologies Industrial Chair in Interactive Technologies, received the CHCCS Achievement award in 2007, and was elected to the ACM CHI Academy in 2005 for his overall contributions to the field of Human Computer Interaction. See www.cpsc.ucalgary.ca/group/lab.