

CONTEXTUAL TESTING IN THE MOBILE PHONE DESIGN PROCESS; A CASE STUDY

A. Woolley, G. Loudon and S. Gill

Keywords: in-context testing, information appliance development

1. Introduction

There has been a growing trend in design research towards the application of research techniques that aim to deliver an insight into a products context, in order to inform product development. Techniques such as Contextual Interviews [Beyer & Holtzblatt, 1998] or Diary Studies (to name but two approaches) that investigate the context a product will be used in, have been shown to deliver insights that can drive innovation in the development of new products and services. For information appliance design, where prototypes can be difficult to generate early in the design process, this early exploration of a products context of use can be difficult to achieve.

The combination of bespoke hardware and software required by interactive products requires prototypes to be created that combine both the digital and physical elements of a product. There have been several prototyping “toolkits” developed to address this early stage prototyping challenge, ranging from low fidelity approaches such as *BOXeS* [Hudson & Mankoff, 2006], to smart toolkits such as *d.Tools* [Hartmann, et al., 2006] that include a wide range of analogue components. Empirical evaluations of the effectiveness of these toolkits has so far focused on laboratory based testing studies [Gill, et al., 2008], and whilst there is no doubt this is an effective evaluation approach, the question arises of whether this is the only evaluation approach that should be supported.

2. Contextual testing with mobile devices in the wider literature

Case study research presents an argument for conducting testing in the field to uncover system requirements by observing natural, unrestricted user behaviours. For example, in studies such as in Tamminen *et al.*'s [2004] study of mobile use in-context, behaviours such as ‘sidestepping’ when using navigation systems would be unlikely to be uncovered in a laboratory test. Mäkelä *et al.*'s [2000] field study of wireless communication using images found that ‘field trials helped ... to understand the possibilities of the product concept much better than any other user research [they had done] before’.

In addition, there have also been empirical studies that have investigated the benefits of using in-context testing as part of a scripted testing approach. The study by Nielsen *et al.* [2006] argues that conducting evaluations in the field uncovers significantly more usability problems than laboratory based studies. In contrast Kaikkonen *et al.*'s [2005] comparative test of laboratory tests and field trials concluded that:

‘when performing a user interface evaluation of mobile applications and devices, field testing may not add significantly to the validity and thoroughness of the test’

This conclusion was based on the number of usability issues uncovered and task successes or failures, contrasted against the amount of time taken to do the tests. However, in the qualitative data from the study, testing in context did have an impact on user behaviour for a download task, where users would divert their attention elsewhere whilst waiting for a download. Kaikkonen *et al.* state that this ‘may

have implications on how the users notice various progress indicators in the application'. So although there was little difference in overall user performance for the download task, Kaikkonen et al. did uncover a difference in behaviour that was relevant to the design of the product when testing in context.

The perspectives presented here from the wider literature highlight two different roles of contextual testing in the design process. Qualitative inquiry work with prototypes aiming to uncover insights to drive innovation, and empirical studies that have largely focused on whether testing in-context provides more accurate usability data by closing the ecological gap when testing.

The case study with a Multinational Mobile phone manufacturer presented in this paper investigates both of these aspects of testing in-context, to highlight the different roles in-context testing with users plays in the design process and the challenges faced when implementing these approaches. The case study investigation was conducted to address the following aims:

- To understand the role of contextual testing in the design process
- To highlight the challenges faced when implementing contextual testing

The study's overall objective of the study was to deliver findings that form a basis for the future development of early stage contextual testing approaches for information appliance design.

3. Investigating the use of contextual testing in a commercial setting

The case study with a Multinational Mobile phone manufacturer involved observation and interview of a usability team consisting of four members; comprising of a senior manager, a team manager and two researchers. The team had a wide range of input on both hardware and software issues, with a slight bias towards software issues as they were physically located in the same office as the User Interface (UI) design team. The usability team as a whole were very experienced in the field, with the least experienced member of the team having worked in this field for over two years.

3.1 Methodology

The case was approached as a single, exemplar case of a large Multinational Mobile phone manufacturer. The overall approach taken has been largely based on that proposed by Stake [1998]. The activities and perspectives of the team were studied using a mixture of unstructured and semi-structured interviews, observation and also documentary evidence. Evolving interpretations were checked with participants to ensure participants perspectives were represented accurately.

A longitudinal approach to the case study was adopted, with 9 visits conducted over the course of a year. Visits were targeted to view different types of user studies, and also to conduct interviews with the team. Using this approach it was possible to see how the team's processes changed and evolved across different projects, presenting a rich picture of their activities. This helped to uncover where processes, opinions and issues applied across projects, which was important as it became apparent over the course of the study that the approach taken for each project was not necessarily the same.

4. Overview of the team's activities

Before addressing the role of contextual testing in the team's activities, it is useful to have a broader knowledge of their overall activities. The usability team studied in this case is small, with four core members. Their primary role is to support the UI designers who are also based on the same site, although they do also have input on hardware issues (such as input devices, types and shape of controls). The team's activities can broadly be broken down into three categories:

- *Planned research*
The umbrella term 'planned research' was applied to any activities where the team planned a formal study, involving either internal or external users to evaluate an element of the design.
- *Quick functional tests*
The team have some involvement in assessing functional aspects of the phone such as the backlight level or which text input engine generates the least errors.
- *Ad Hoc consultancy*

The term ‘Ad Hoc’ consultancy covers instances where the team are engaged with the UI designers in a consultancy role. For example, a member of the UI design team might seek advice from a usability team member on an issue and this might then be resolved by a quick five minute conversation.

The team’s time is relatively evenly split across all three of these tasks, bearing in mind that user interviews take a lot of preparation and analysis time. It is important to understand that Ad Hoc consultancy pervades a lot of the teams work, for instance they might be planning the next study, working on the report for the previous study and involved in Ad Hoc consultancy all in the same day.

4.1 Role in the design process

Figure 1 represents a high level view of where the usability team’s activities fit into the overall design processes of the company. The team in this case study are somewhat detached from the main ‘Concept Generation’ work that goes on in the overall processes of the main company, with the majority of their time taken up by ‘Design Iteration’. The team do occasionally input into the concept generation process in the form of a concept usability report which details any issues that might be foreseen from a usability perspective. The team are well integrated into the UI design process; however, with respect to hardware they tended to receive prototypes late in the design process and they had less input into this element of the design.

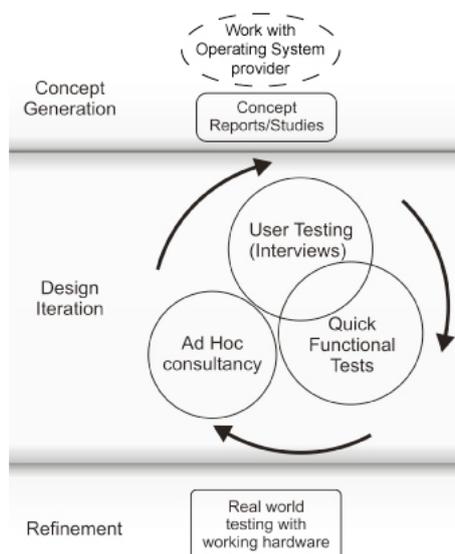


Figure 1. Team’s activities in relation to the design process

4.2 A Toolkit approach to usability

Different members of the team discussed on several occasions that they were aiming to develop a “toolkit” of research methods that they would be able to select from, in order to give an appropriate response to the challenges presented by each product over the course of its development. There were a number of reasons given for this toolkit approach to the application of research methods in the design process. Firstly, each project presents different challenges and the team need to have a process agile enough to cope with these changing demands. Secondly, mobile phones are increasingly complex devices, and therefore the team need to be able to direct their activities towards areas where issues are most likely to occur.

In terms of the research methods used, planned research was largely conducted using an interview based approach, often with users running through predefined tasks with a prototype as a part of this. However, as a part of this ‘toolkit’ approach, the team did use other approaches such as the use of subjective rating scales depending on the demands of the issue they were investigating.

4.3 Decisions made early in the process

An important factor in the team's processes was that decisions that have a major impact on the design of the phone were often taken very early on in the design process. With most of the phone projects there was a legacy from previous interfaces and applications, and also conventions laid down in the phones operating system (the phones that the team work with were based on an operating system developed by another company). Therefore the product was rarely being designed from a blank slate, and there were occasions where design recommendations made by the usability team could not be implemented because of restrictions laid down in the operating system before the project had started. Therefore, in terms of the length of the project timeline for each phone under development it is important to see this as extending well before the handset design actually starts.

Hardware decisions in particular were often made very early on in the design process relative to the usability teams input. The team was often seeing prototypes of the hardware very late on in the process, although this was an element of their processes that was beginning to change, with block models of the phones form becoming available to the team earlier than had previously been the case. However, the team were still finding that they were conducting most of their hardware testing very late in the design process, and this was seriously limiting the design changes that could be made on the basis of usability studies recommendations.

"I know that the (hardware) changes that you can make ... they're just like minor changes so bearing that in mind we've always suggested changes that we think might happen rather than like you need to completely change this button because that'll never happen. But if we just say can you raise it slightly so people can press it more easily..." - Researcher

5. Where contextual testing features in these activities

Within the team's testing activities context featured in a number of areas. Although previously qualitative, upfront exploratory studies had been conducted (for example investigating users habits with existing music players), a recent change in project structure with more handsets being designed in parallel meant that time pressures had increased such that the team felt they no longer had time to conduct this sort of purely qualitative contextual inquiry.

5.1 No usability laboratory

Despite this pure contextual inquiry work not being part of the team's processes, information about a products context did feature in their processes. The team do not have a usability laboratory set up in the traditional sense and instead often visited users in their own context to conduct user studies. The team had previously used a laboratory set up for testing but had abandoned this approach in favour of a more mobile one. The reasons given by the team for this were a mixture of both practicality and methodology. This account was given by the team member responsible for instigating this change in approach:

"To get ... phone users to come to us is pretty tricky. I mean, they might be earning quite a lot of money, so to drive ... to us to do something was a bit tricky. We then put them into a completely new environment. I felt it was quite useful if somebody said what happens with my pc is this, and when I put my Nokia in it does that. I'd much rather people showed me in context and try to understand the context. They're more relaxed in their own environments." - Senior manager

Therefore in addition to the methodological benefits of setting the test in context, the practical advantages of being able to travel to users were an additional driver behind this adoption of a mobile approach. The element of time was something that was key for the team, and was also cited as a reason why the team had reduced their use of video recording in studies. Faster reporting and iteration of designs was seen as more important than taking longer to go through video to capture absolutely all of the issues possible.

"I tend to think if we take good notes and have a good process for observing and taking notes, that we pick up 95% maybe 90% of what we need to know. I'd rather do that, feed that back to the designers quickly, [and] get them to redesign ... rather than spending lots of time going through video." - Senior manager

As a result of not using a laboratory set up, the team were often visiting users in their own context, whether this was a place of work or in someone's home. The team all expressed the opinion that users

were more relaxed in their own context although this was often expressed in relation to the ‘feel’ of the test, rather than having an actual affect on the outcome of an evaluation.

Where seeing users in their context was thought to make a difference, was in allowing the team to gather wider information about users and their context. This somewhat anecdotal collection of additional information was something that was a feature of almost all tests observed that contained some element of interview. Whether this information was fed back into the design process largely depended on the nature of the study. For more functional evaluation work, it was less likely that these additional comments would be fed back than for an early stage concept study.

5.2 Testing in the real context of use as a quality process

The main area of the design process where testing the product in its real context of use was a major part of the process was quality testing. This occurred towards the end of the product development cycle, generally once the product had passed the ‘Alpha’ milestone (code complete and functioning but still in a prototype state) and used working hardware and software prototypes. Although the team were not running this testing directly, they were responsible for handling the feedback related to usability issues. These real world testing activities included:

- Use of handsets internally by employees
- Alpha and Beta testing with both internal and external users
- An additional testing programme that ran from Alpha until the end of the project that included testing of software and accessories such as Bluetooth peripherals alongside the testing of handsets

The team member responsible for handling the feedback from this testing pointed out that once the phones start being used realistically, reports of problems and requests for changes “balloon”:

“You don’t find problems unless you use the phone in real world situations. It’s all very well sending one text a day but until I’ve sent say ten text messages and received thirty or forty and filled the inbox, I’m not going to come across real world problems. You have to use it realistically”- Team manager

Phones are used within the confines of the office by employees up to the point when the phone has officially been released, at which point they are able to take it outside the office and use the handset more realistically.

Testing in the real context of use was very much limited to the final stages of the design process both by issues of confidentiality and the availability of functional prototypes. Therefore, this testing was conducted as a way of refining the design and building quality into the software (preventing crashes etc). Although recommendations for changes could be raised as a result of these tests, because the project was so far advanced these tended to be relatively minor and any major changes would have to be carried over to the next project. Availability of prototypes and the level of software development (i.e. developed to a level that people can reliably use the product as a mobile phone) were the main factors that limited this testing to a later stage activity.

5.3 Previous experiments with in-context testing earlier in the design process

The team recounted two instances where they had trialled different user testing methods that included contextual elements. As a whole the team were very open to trialling new approaches and expand their toolkit of methods where possible. The two instances discussed here are; leaving early prototypes of hardware and software with users for an extended period; and also conducting a study on screen glare with the user walking.

5.3.1 Leaving early prototypes with users

The team had previously trialled leaving early, somewhat unstable prototypes of working hardware and software with users for several weeks (previously, only later stage, more stable releases had been used for this style of testing). The decision was made to leave the hardware with the users for more than a week, as it was thought unlikely they would realistically spend time synchronising contacts with the PC, using calendar functions etc if users only had hardware for a short period.

It was the initial intention of the team to hold a weekly phone conference with users to discuss issues. However, in reality it turned out that much more contact time was needed, with users perceiving the usability team members as on hand at any point to help with any queries or problems. One user also had significant problems with software crashing, and a lot of time was spent investigating this before it was discovered that this issue was caused by additional third party software the user had loaded on to the prototype themselves. This issue was caused by the user's perception that the prototype was in fact a fully functional product.

The opinion of the usability team was that on reflection, although this activity was worthwhile and useful, the process would need to be much better managed in terms of commitment of resources, particularly time, if they were to conduct it again.

5.3.2 Testing whilst walking

One member of the team recounted the following experience of piloting a usability test that situated the user in-context in order to include different modes of interaction:

"I did one study and I tried to get people to walk whilst they were doing it because it was to do with reflections on the screen. ... I wanted to do it in natural settings so people walking outside, trying to use their phone and multitask ... it just wasn't possible because for me to run along side them with my pen and paper and try to see what they were doing, it just wasn't doable really..." - Researcher

When asked how users had reacted to this approach, the team member responded:

"They felt stupid! That was the main reason I didn't do it. It doesn't matter if I'm running alongside them really, but if they're feeling a bit silly..." - Researcher

As a result of this negative user reaction to the pilot, the usability team member chose to eliminate the element of walking from this study. Above all the priority from the team was to make sure users were comfortable with testing methods used, and they felt that this approach was compromising this.

6. Reflections on contextual testing

The team didn't necessarily feel that their processes as they stand at the moment were missing anything as a result of not including pure contextual inquiry work (say, observing users' interactions with various music players). They speculated as to whether this was somewhat outside their present remit, and may be within the remit of other departments in the company. For example:

"I think it would be useful for product planning, the people who decide what goes into the phones or how things are going to be" - Researcher

However, the usability team is part of a constantly changing set of organisational processes, and over the course of the case study, the team were becoming involved earlier in the design process and beginning to conduct more concept studies for departments such as product planning. In spite of this viewpoint that contextual inquiry work may fall outside of their current role, it was an area members of the team expressed a personal interest in doing more work in.

An underlying reason given for the interest in conducting the more qualitative studies given by two members of the team was to increase the knowledge within the team in relation to mobile phone design. Not on a project specific basis, but as a way to build up a pool of knowledge to direct other work.

"I think it would be great because it would give us more of an understanding of what's important to people ... where to focus our resources and give us clues as to what's good in design. I think again it's the ideal really, that would be great just to increase our awareness our understanding that kind of thing. I don't think it's entirely necessary for the job and the scope of our job but it would be good." - Researcher

This is an interesting viewpoint on contextual testing, and one that fits with the perspective that contextual testing can be used to gain insights that can direct future development. What is of particular interest here is how this knowledge could be applied across a range of projects, rather than being specific to only one project.

An example given by another team member that relates to this issue describes where a team member negotiated dedicated hardware volume keys at a very early stage:

“So I supposed it was based a bit on logic, a bit on ... I remember doing some user studies a long time ago where users expected separate volume keys.” - Senior manager

Although this team member (who had actually conducted the studies) played down the significance of this previous study, a different team member counted it as being of much more importance in making the decision to have separate volume keys on the product. This element of negotiation on the basis of previous studies is important, as the team were often engaged in expert analysis of products as part of Ad Hoc consultancy activities, in situations where resources weren't available to conduct formal usability studies.

7. Discussion

The process seen in this company represents a pragmatic and flexible approach to usability, and this process maps well to other published work, such as the approach of Nokia detailed by Lindholm *et al.* [2003]. The contribution of contextual testing seen across projects, and its support of a range of the teams activities, is important. With the majority of published information appliances in the literature presenting a linear view of the design process, this cross project contribution is important to consider. Processes presented in the published literature are by necessity simplifications of the design process, and it is not the intention to criticise this fact here, rather to reinforce the point that when considering the return on investment of contextual testing, it is important a long term view is taken.

Leaving even late stage prototypes that were relatively functional with users was seen to be an intensive exercise. Particularly as users had been unclear as to the limitations of the prototype. Given that decisions that were difficult to change, particularly related to hardware, were made so early in the design process, if contextual testing is delayed to a stage where functional hardware is reliable, it is likely that conclusions would be severely restricted. However, there is a tension here, as testing in context was a key quality measure and by testing applications in real world usage scenarios, the number of issues uncovered was seen to 'balloon'.

Earlier stage testing approaches such as that used by Mäkelä *et al.* [2000], move contextual testing to an earlier stage in the design process by limiting the functionality that is implemented in a product prototype (in this case, limiting functionality to communication with images only). The team had commented that they tended to focus efforts on new or novel applications, and the approach proposed by Mäkelä *et al.* would certainly work well with this philosophy. However, the approach described by Mäkelä *et al.* still requires significant development resources to implement and time to support. The early stage prototyping tools such as *BOXeS*, *d.tools* and the *IE system* [Gill, et al., 2008] present an opportunity to develop interactive prototypes much faster and at an earlier stage in the design process. However, the low levels of functionality implemented by these systems would need to be accounted for when implementing in-context testing, as it is clear that even with highly functioning prototypes, unrestricted in-context testing requires high levels of support.

8. Conclusions

This case study has highlighted some of the challenges faced when implementing testing in context. As decisions, particularly related to the design of hardware, are taken early in the process, if contextual testing is to have an impact on the final design of the product, it must be conducted early in the design process. Contextual testing was seen to be an important quality measure in the development of products; however, it was difficult to integrate this approach early in the design process due to functional prototypes only being available at a very late stage.

To add maximum value to the design process, contextual testing methods should ideally be implemented early in the design process, using rapid and flexible tools that support a 'toolbox' approach to usability evaluation. Information appliance prototyping toolkits present a possible solution to prototyping challenges at this early stage, but it is important to investigate the application of these tools as a part of contextual testing approaches, particularly given the limited functionality available using these techniques.

If contextual testing can be implemented early in the design process using rapid, low investment techniques, it is likely that this would have long term benefits to the design process. If early in-context

testing can deliver findings that add value across projects, the return on investment may justify more complex and time consuming implementations of this approach.

9. Future work

It is important that low investment, early stage information appliance prototyping toolkits are evaluated in relation to in-context testing environments. This is needed order to establish both if this is a viable approach to in-context testing and also to establish the limitations on the data gathered using such early stage prototyping tools that have limited levels of functionality.

Acknowledgements

The authors would like to thank the usability team in this study for their kind co-operation with this research.

References

- Beyer, H., & Holtzblatt, B. (1998). *Contextual Design: Defining Customer-Centred Systems*. San Francisco: Morgan Kaufmann.
- Gill, S., Walker, D., Loudon, G., Dix, A., Woolley, A., Ramduny-Ellis, D., et al. (2008). *Rapid Development of Tangible Interactive Appliances: Achieving the Fidelity / Time balance*. *Tangible and Embedded Interaction, a special issue of the International Journal of Arts and Technology*, Vol. 1, No. 3/4, 2008.
- Hartmann, B., Klemmer, S. R., Bernstein, M., Abdula, L., Burr, B., Robinson-Mosher, A., et al. (2006). *Reflective physical prototyping through integrated design, test, and analysis*. *Proceedings of UIST 2006*. Montreux, Switzerland.
- Hudson, S. E., & Mankoff, J. (2006). *Rapid construction of Functioning Physical Interfaces from Cardboard, Thumtacks, Tin Foil and Masking Tape*. *Proceedings of UIST'06* (pp. 289-297). Montreux, Switzerland: ACM.
- Kaikkonen, A., Kallio, T., Kelalainen, A., & Cankar, M. (2005). *Usability testing of mobile applications: a comparison between laboratory and field testing*. *Journal of Usability Studies*, 1(1) , p.4-16.
- Lindholm, C., Keinonen, T., & Kiljander, H. (2003). *Mobile Usability: How Nokia Changed the Face of the Mobile Phone*. McGraw-Hill.
- Mäkelä, A., Giller, V., Tscheligi, M., & Sefelin, R. (2000). *Joking, storytelling, artsharing, expressing affection: A field trial of how children and their social network communicate with digital images in leisure time*. *CHI'2000* (pp. 548-555). ACM Press.
- Nielsen, C., Overgaard, M., Pedersen, M., Stage, J., & Stenild, S. (2006). *It's Worth the Hassle! The Added Value of Evaluating the Usability of Mobile Systems in the Field*. *Proceedings of NordiCHI '06: Changing roles* (pp. 272-280). Oslo: ACM.
- Stake, R. (1998). *Case Studies*. In N. K. Denzin, & Y. S. Lincoln, *Strategies of qualitative inquiry*. London: SAGE.
- Tamminen, S., Oulasvirta, A., Toiskallio, K., & Kankainen, A. (2004). *Understanding mobile contexts*. *Personal Ubiquitous Computing* , 8: 135-143.

Dr Alex Woolley
National Centre for Product Design & Development Research,
University of Wales Institute, Cardiff, UK
T: 02920 41 7009
E: alex.woolley@gmail.com