TOWARDS DESIGNING MORE EFFECTIVE SYSTEMS BY UNDERSTANDING USER EXPERIENCES

By

Clare J. Hooper

A thesis submitted for the degree of Doctor of Engineering

School of Electronics and Computer Science,

University of Southampton,

United Kingdom.

December, 2010

UNIVERSITY OF SOUTHAMPTON

ABSTRACT

FACULTY OF ENGINEERING

SCHOOL OF ELECTRONICS AND COMPUTER SCIENCE

Doctor of Engineering

TOWARDS DESIGNING MORE EFFECTIVE SYSTEMS BY UNDERSTANDING USER EXPERIENCES

by Clare J. Hooper

This thesis is about social technologies, user experiences and the problems of creative design. It is motivated by a desire to give people who are offline the access to social technologies that is currently provided via the web. There exist technologically-oriented approaches to solving this problem, but their focus on technology comes at a cost of neglecting the experiential aspects which motivate the work. This focus can result in systems which are functional but unappealing to (or even unusable by) their target audiences.

After describing the motivation for the work, this thesis explains the state of the art and presents an exemplar system built with a technological focus. This thesis then presents Teasing Apart, Piecing Together (TAPT), a Software Engineering design process developed to address this gap in the field of software design. TAPT enables the understanding of user *experiences* and scaffolds the redesign of these for new contexts.

After explaining the TAPT process and how it was built, a three-phase mixed methods evaluation is described. This consists of a large-scale comparative evaluation, an expert review of the outputs of that evaluation and case studies grounded in industrial and academic practice. The results of these evaluations show that TAPT, which can be used in an agile manner, provides a strong analytical framework for understanding experiences and supports the redesign of experiences in new contexts.

Contents

List of Figures

List of Tables

Acknowledgements

I'd like to thank Dave Millard for his support and guidance through the course of this research. His uncanny ability to ask searching, difficult questions has undoubtedly strengthened my critical abilities, not to mention this work. My thanks also extend to Andy Stanford-Clark, whose shrewd input and enthusiasm have kept me on track and engaged.

I've had the good fortune to have guidance from many others, notably including my second supervisors Dave de Roure and Mark Weal. I'm grateful to the many study participants who so generously gave up their time, and for the discerning input of my brilliant proof readers. This work was also made possible by IBM's generous support.

The communities at IBM Hursley and LSL have been welcoming and supportive, while figures such as Faith Lawrence and Max Wilson been role models for conducting doctoral-level research. I continue to draw inspiration from those around me, particularly Wendy Hall and Nate Matias, who achieve great goals with dedication and hard work while supporting the many communities around them.

I'm lucky to have many wonderful people in my life, and I'd like to thank them for their support: amongst people too numerous to list, my thanks extend to Reena, Jane, Don, Thomas, John-Mark, Ruth 'n' Tim, Alan, Hugh, Dave and the OGC crowd. Thanks for putting up with me!

Finally, and most of all, my thanks to Alisdair for his unfailing support, which is too extensive to enumerate.

Chapter 1 Introduction

This thesis is about social technologies, user experiences and the problems of creative design. The motivation behind this work lies in a desire to give people who are offline the access to social experiences that is currently provided via web-based technologies. There exist technologically-oriented approaches to solving this problem, but they involve a focus on technology instead of experiential aspects. This focus on implementation at the cost of understanding user experience can result in systems that are functional but unappealing to (or even unusable by) their target audiences.

In response to this problem, the author built a design process called Teasing Apart, Piecing Together (TAPT). TAPT is a two-phase approach that involves deconstructing an experience in order to better understand it, and then rebuilding it in a new context. It has broad applicability as a tool for creative design and analysis, as well as being a potential solution to the problem described above. This thesis presents existing work in the area of creative design processes, describes the process by which TAPT was conceived, and recounts a mixed methods evaluation of TAPT.

1.1 Motivation

The work towards understanding user experiences presented in this thesis is motivated by a desire to enable broader access to social technologies. Such technologies include more traditional tools such as email and landline phones, fully established technologies such as mobile phones and instant messaging (IM), and newer examples including social networking sites and blogs. These media can be a source of fun (Hart, 2008) and emotional support (Wright, 2002), yet a large chunk of society is excluded from this arena of social interaction. For example, in general, elderly people face a range of obstacles to the uptake of technology (Namazi, 2003).

Existing work has demonstrated that the elderly benefit from being online (Cody, 1999) (Nahm, 2001) and on social networks (Colvin, 2004) (Morris, 2005), and includes examples of social networks for the elderly (Waterworth, 2009).

The issue of exclusion from online social experiences is increasingly important, especially given the trend for ageing populations across the western world, and the geographical dispersion faced by many families. Technologies built without due consideration can exclude users considered ‘non-standard’, whether due to impairments, economic considerations or cultural aspects. Meeting these considerations when building any system requires a holistic outlook and an inclusive approach to design (Newell, 2004).

If it is possible to connect offline people through technology with which they are familiar, their general wellbeing could be improved. For example, contact with family might become easier (children could email grandparents regardless of whether the grandparents own a PC), and access to online content could be gained without requiring ownership of expensive computers or mobile phones, and knowledge of how to use web browsers. Many more people would be able to access the online content and communications facilities that so many of us take for granted. Consider the following scenario (Hooper Owens, 2009b):

> *Gerald is an elderly man, who lives alone in a flat. He has no interest in computers, but is happy to benefit from services enabled by technology. For example, he is very interested in seeing photographs from his grandchildren's sports day. A digital photo frame is installed on his wall; when Gerald’s son Matt uploads photographs from the sports day to a social networking website, the pictures are shown in turn on the display.*
>
> *Later that day, Gerald's grandson Billy writes him an email about the sports day. Gerald doesn't access his email account on a PC. Instead, the email is translated into an audio file and read to him via his telephone; an alternative way, which Gerald does not prefer, would be a paper copy via a microprinter. Gerald listens to Billy's message late in the day, when Billy is probably asleep. Gerald records a reply, a voice message for Billy to receive in the morning.*

Gerald is also particularly interested in a small community of ex-pat friends of his, who live in Spain. He plans to travel to visit them one day, but meanwhile he follows updates about their day-to-day lives online. Gerald accesses these updates via a teletext[1]-style display on his television.

In order to effectively build the social aspects of such a system, it is necessary to translate the web-based experiences offered by social sites into a new modality. Gaining a deep understanding of the experiences offered by such websites is a vital step in allowing them to be re-imagined in a new context.

Prior work has resulted in a plethora of pervasive systems in research labs, and some of these systems are socially-oriented. Existing work also includes discussions of the issues of accessible interfaces for novel devices and the ethical issues of working with vulnerable users. Other socially-oriented work considers virtual communities, devices to facilitate face-to-face social interactions and devices to increase a sense of presence for geographically disparate groups.

There also exists prior work in the area of understanding user experience and supporting software design: creative software design methods are part of the field of Computer Science, and although traditional methods do not incorporate experiential aspects, emerging approaches begin to address this area. There exist various models of experience, and approaches to understanding user experiences. Related to this area is the existence of frameworks for evaluating creative methods.

Although this prior work paves the way, there is a lack of methods for understanding and *replicating* experiences: existing work does not fit together in a way that provides a suitable approach. When attempting to understand experiences, some aspects (particularly social and emotional factors) are not well-served by existing methods.

For example, existing tools and techniques support the provision of pervasive social technologies that could be framed as alternatives to current web-based social networking. Such approaches include the creation of a list of functional requirements and building a system to meet these requirements, perhaps even guided by user-oriented methods such as Scenarios and Personas. Use of Scenarios involves

---

[1] Teletext is a text-based television information retrieval service, which runs in the UK.

identifying characters (called personas) that are representative of user groups who will use the future system, and writing scenarios of how those personas would use the envisioned system.

Although the above approach could result in a technologically-functional system that might be built in a way which acknowledges the diversity and needs of end user-groups, if it was an attempt to re-provide existing experiences in a new context the approach would have failed to fully examine those experiences. The new system might well replicate aspects of the original experiences, but those aspects would be highly likely to be connected with the original's superficial design (which is obvious to see) and functional aspects (which are measured with the list of functional requirements). The resultant system may or may not be successful, but it almost certainly would fail to fully address the experiential aspects of the original system, simply because no mechanism was in place to attempt to understand or replicate these aspects.

Existing methods to attempt to understand user experiences could be applied in this arena, yet they do not guarantee coverage of all relevant aspects, and are not constructed in such a way as to facilitate the redesign of existing systems for new contexts.

TAPT, the method presented in this thesis, is intended to solve this problem: it is a process that involves systematically analysing a user experience towards a) understanding and b) replicating that experience. As is demonstrated in Chapter 4, TAPT enables the identification of experiential aspects of interactions, and in particular supports the identification of abstract and emotional aspects.

By improving their understanding of user experiences, designers are better able to create appropriate interfaces for unfamiliar groups of users. In her keynote presentation at the 2010 Web Science conference, Gilbert (Gilbert, 2010) described the digital divide and the vastly different frameworks of ICT use which exist, contrasting for example ICT use by poor women as opposed to employed professionals. She remarked that we cannot build a better web from our privileged technology-use framework, and spoke about the necessity of understanding the experiences of the people for whom we build technology.

Better understanding of user experiences also improves designs in a broader context, enabling designers to build systems that incorporate familiar aspects. People are comfortable with familiar experiences, which require no learning curve and feel straightforward.

Technology-rich interactions can resemble low-fidelity interactions and vice versa: Nielson reports that at the Hypertext 1987 conference Ted Nelson described Apple Macintosh computers as “elaborate paper simulators” (Nielsen, 1988), while recent work involves augmenting physical paper interfaces with computer-like facets (Laio, 2010).

This thesis details research conducted in response to the above issues. Inspired by the concept of ‘experience deconstruction’ (Dix, 2003) the TAPT process was developed to support Software Engineers in analysing and redesigning experiences (see Chapter 4). A three-pronged evaluation of TAPT was conducted:

1) A comparative evaluation, in which 43 Software Engineering professionals tackled design tasks using TAPT or one of two other methods (Chapter 5).
2) An expert review, in which six domain experts evaluated the design artefacts produced in the comparative evaluation (Chapter 6).
3) Four case studies, in which participants from industry and academia applied TAPT freely to their real-world work problems, and gave feedback on its usefulness (Chapter 7).

1.2 Hypothesis and Research Questions

This thesis addresses the following hypothesis:

A systematic process of user experience deconstruction (that identifies superficial, abstract and emotional elements) can improve the critical analysis of user experiences, and in a software development context better scaffold the initial design process or evaluation of user experiences compared to existing methods.

A ‘systematic process’ is a repeatable, step-by-step procedure, which can be reliably reused by a variety of people in different contexts.

‘Critical analysis’ refers to the scrutiny of user experiences towards some analytical end, for example to gain insight into user experiences across a set of similar pieces of software.

‘To scaffold’ refers to the provision of support to facilitate the creation of initial design ideas for a software system. That is, by using a deconstructive method, practitioners find themselves better able to conduct initial design work, particularly by producing better visions of their final systems.

'Evaluation' refers to an act of comparative evaluation. For example, practitioners may analyse multiple user experiences (for example, perhaps considering experiences before and after a system has been built) and compare those experiences.

This thesis presents an evaluation of TAPT towards answering the above hypothesis. A number of research questions were formulated to break the hypothesis down:

1. Is the TAPT process of user experience deconstruction systematic?

It is important to ascertain whether TAPT is systematic, because if it is not then it cannot be used in practice. This question can be answered by investigating whether participants in experiments are able to apply TAPT, and also by gauging participants' views on the systematic nature of TAPT.

Chapter 5 and Chapter 7 describe TAPT's use by participants, and Chapter 5 also discusses participants' opinions on TAPT's suitability for workplace use and how structured it is.

2. Can a TAPT analysis identify superficial, abstract and emotional elements?

The identification of superficial design elements is a relatively straightforward task, as by nature these elements are not subtle or difficult to observe. The identification of such aspects was included in the TAPT process in order to ensure practitioner coverage and understanding of these aspects of the experience: it is important to test whether TAPT supports this identification.

It is particularly important to understand whether TAPT facilitates the identification of abstract and emotional elements, which are much less obvious to the observer. It is the difficulty of understanding these aspects of experience that drove the creation of TAPT.

Mixed methods provide a route to understanding whether TAPT achieves these objectives: insight can be gained from numeric participant ratings and qualitative comments. The evaluative work described in Chapter 5, Chapter 6 and Chapter 7 includes results relevant to this question.

3. Can a TAPT analysis be used to scaffold the initial software design process?

This is a question about TAPT's applicability to a specific problem domain, that of designing software. TAPT was developed to support this process, so this question is particularly important. Again, insights can be gained through mixed methods, including laboratory-based experiments and case studies of TAPT's use for this purpose. All three experiments (described in Chapter 5, Chapter 6 and Chapter 7) provide insights into this area.

4. Can a TAPT analysis be used to evaluate user experiences to support software development?

This is another question about *how* TAPT can be used, in this case in an evaluative sense: a software system can be critiqued by comparing analysis of user experiences with and without the system, or by comparing experiences between the system and a rival.

The primary goal of TAPT is to scaffold software design, but if it can demonstrate value in analysing user experiences in other ways then this is a positive result. Although laboratory-based studies could be conducted to provide information in response to this question, these would represent a rather heavyweight approach to answering a question which is not the primary focus of this work. As such, a lightweight approach such as case studies (in which practitioners attempt to use TAPT towards such an end and report on its efficacy) seems appropriate. Results in Chapter 7 address this question.

5. Can a TAPT analysis be used for critical analysis of user experiences?

This final question also concerns how TAPT can be used in senses other than for software design, focusing instead on understanding user experiences towards an analytical end aside from the evaluation of software.

As discussed, TAPT's primary purpose is to scaffold software design and so the studies described in this thesis do not focus strongly on this particular question. However, a case study in Chapter 7 provides evidence about TAPT's effectiveness when used in this manner.

In summary, this thesis aims to answer questions about TAPT's suitability for systematically identifying superficial, abstract and emotional effects within experiences, and about TAPT's applicability in the realms of software design, software evaluation and critical analysis.

1.3 Contributions

This thesis describes TAPT and shows that it is a structured method for thoroughly exploring facets of a given experience, including those emotional and social effects that might otherwise be overlooked: identifying these effects is essential for understanding an experience in an abstract sense so it can be deconstructed. Current tools and techniques do not appear to support the understanding of these aspects of experience.

The aim of this work is to identify the strengths and weaknesses of TAPT and identify potential improvements through a comparative evaluation, expert review and in-depth case studies.

This thesis documents three key contributions made to the field of user experience:

1. The TAPT method, described in Chapter 4, including the approach, a systematic process for applying it, simple structures for recording the results, and examples of its use in different contexts.
2. A three-part mixed methods analysis of TAPT, presented in Chapter 5, Chapter 6 and Chapter 7. This consists of a comparative evaluation, expert review and in-depth case studies of TAPT's use in industry and academia, which provide information on TAPT's strengths and weaknesses and the ways in which it can be applied.
3. Reflections on approaches to analysing Software Engineering design processes such as TAPT: these are given within each evaluative chapter while an overall, holistic reflection is in Chapter 8.

1.4 Scope

Horváth suggests there are three types of research within industrial design (Horváth, 2007, Horváth, 2008):

1. Research in a design context: analytical disciplinary research, aiming at insights, understanding, and predictions. It relies mainly on the knowledge and methods of background disciplines, and is theoretical.
2. Design inclusive research: this blends research and design, focused on contextualization and integration.
3. Practice-based research uses research-inspired principles in practice to answer questions that emerge from practice. It is conducted to gain intelligence for the practice, rather than insight for design science.

The work presented in this thesis primarily constitutes research in a design context, although use of case studies involving real-world design use of TAPT situate the work slightly closer to design inclusive research than otherwise.

The results of this research are primarily sourced from the domain of Software Engineering in the UK, although a small portion of the data concerns academic research in the UK and Norway. Although Software Engineering processes are similar in other countries, the results must be considered in this cultural context.

1.5 Document Structure

This thesis describes the motivation for TAPT, a method for analysing and redesigning user experiences, and its development and evaluation.

Chapter 2 provides material relevant to the motivation for this work, presenting material on pervasive technology, issues of accessibility and ageing, and social technology. It then provides further evidence for the motivation by describing an example of the problem: it presents a sample social infrastructure and its evaluation, and explains how a better understanding of experiences was required in order to have confidence in interfaces built on top of the infrastructure.

Chapter 3 discusses prior work in this area, first discussing creativity and its place in Computer Science, then moving on to software design. It next discusses User Experience, describing the emergence of the field, models of experience, and methods for understanding experience. The chapter closes with a discussion of the evaluation of creative design methods.

Chapter 4 presents TAPT. It first describes ‘experience deconstruction’ as presented by Alan Dix (Dix, 2003), before explaining the steps taken to formalise the concept of deconstruction into the TAPT process. Chapter 4 also presents the results of an initial application of TAPT to social networking functionality and reflections on TAPT’s efficacy.

Chapter 5 describes the first of the three experiments conducted to evaluate TAPT, a comparative evaluation. In this experiment, Software Engineers were asked to apply TAPT and other design methods to certain tasks. After outlining initial pilot studies, the chapter describes the experimental design and provides an analysis of the resultant design artefacts and participants' comments on their experiences conducting the design tasks.

Chapter 6 describes the second experiment towards evaluating TAPT, the expert review. Here, six experts were individually presented with blind design artefacts (produced using a number of methods including TAPT) and TAPT artefacts from the comparative evaluation and asked to provide their opinions on them. This chapter explains the experimental design and factors which impacted results, before discussing results in three broad sections: results about blind artefacts, results about TAPT artefacts, and broader insights.

Chapter 7 presents the final facet of TAPT's evaluation, the use of four case studies. Participants, who were Software Engineers and academic researchers, were equipped with the TAPT method and asked to use it in whatever way seemed most appropriate for supporting their work. Chapter 7 first describes the experimental design and each of the four individual case studies. The chapter then discusses participant motivations and expectations, the way in which TAPT was used and comments on TAPT itself.

Finally, Chapter 8 presents overall conclusions. It discusses the overall results presented in this thesis, and in particular provides reflections on the overall method of evaluation. It revisits the hypothesis and contributions before discussing directions for future research.

1.6 Declaration

This thesis describes the research undertaken by the author while working within a collaborative research environment. This report documents the original work of the author except where stated. All trademarks are acknowledged.

Chapter 2 Background and Motivation

This chapter introduces the problem area in more detail, first presenting literature on relevant domain areas (pervasive technologies, issues of ageing and accessibility, and social technologies) and then presenting a social infrastructure which exemplifies the problem: existing tools and techniques support practitioners in addressing technological requirements but fail to assist in understanding experience.

Pervasive (or ubiquitous) computing concerns the availability of many effectively invisible computers throughout the physical environment, wherein their interfaces are so transparent to use that they are effectively invisible to the people using them. This chapter particularly concerns computing in non-work contexts, such as domestic and care environments. Pervasive computing is relevant to the problems addressed in this thesis because it offers a way to introduce functionality which is currently web-based to people who are not traditional web users, via lightweight and relatively cheap mechanisms.

Issues of ageing and design for the elderly are central to the problem outlined in Section 1.1. The accessibility and appropriateness of systems' interfaces is particularly relevant in the context of novel, pervasive interfaces because there are fewer existing usability paradigms when considering new technologies. Additionally, this work focuses on audiences for whom technology may be novel and who may be subject to various disabilities, making accessibility particularly important. Ethical issues are also discussed.

Next, social technologies are considered: these are technologies used to support existing social relationships or facilitate the formation of new such ties. This chapter describes various social systems, with a focus on the functionality and history of social networking websites. An understanding of this area is necessary in order to consider how to improve access to social technologies.

Finally, this chapter describes a novel social system built by the author as an exploration of issues in the realm of multimodal messaging infrastructures. A technologically-oriented approach was taken to building the system, which presents further evidence for the need for better design methodologies, particularly when user experience is a key driver and the context of implementation is novel.

2.1 Pervasive Technologies

2.1.1 Introduction

This section introduces pervasive technologies. Pervasive devices offer a means of implementing the vision described in Section 1.1, where digital technology is integrated seamlessly into the physical environment.

The term 'pervasive computing' appears to have connotations with computer systems (e.g. Varshney (Varshney, 2003), Lorincz (Lorincz, 2004) and Stanford (Stanford, 2002)), while 'ubiquitous computing' seems to have been associated with Weiser's vision of 'calm computing' (Weiser, 1989) (Brown, 1999). That said, the current literature (and therefore this document) largely uses the terms 'pervasive' and 'ubiquitous' interchangeably. For example, Ark and Selker (Ark, 1999), in their introduction to the 1999 IBM Systems Journal, explicitly state "The terms *pervasive computing* and *ubiquitous computing* are used interchangeably throughout this issue." Similarly, Korhonen and Bardram (Korhonen, 2004), in their introduction to the section on pervasive healthcare in the IEEE Transactions on Information Technology in Biomedicine, refer to "pervasive computing—or ubiquitous computing, proactive computing, ambient intelligence."

Pervasive or ubiquitous computing involves the availability of many effectively invisible computers throughout the physical environment (Weiser, 1989): that is, the technologies are so transparent to use that people do not notice them. Weiser suggests this invisibility as analogous to text: people do not explicitly notice or struggle with text in newspapers, books, adverts and on food wrappings, but simply read it if they so desire. Weiser describes pervasive technologies as being the opposite of Virtual Reality technologies, which work to simulate an alternative world: by contrast, pervasive technologies invisibly enhance the existing environment.

Weiser (Weiser, 1993) suggests that traditional computers are in the way of work to be done, not due to their interfaces, but because they demand the focus of those using them. Weiser suggests that pervasive systems can be used by those in shared situations, regardless of their technological skills: people can simply pick up a notebook-sized computer, which is not associated with one specific person but is analogous to (and as easy to use as) scrap paper, and use it, whether in a shared context or not.

The aim of this is to produce ‘calm’ computing, where the technology is not the focus of people’s attention, and where the people using it control the technology, rather than being driven by it (Brown, 1999).

The evaluation of novel interfaces is not necessarily straightforward. Researchers have noted an absence of suitable evaluation tools for mobile and pervasive interfaces (Leichtenstern, 2008) as well for as novel exhibits (Gaver, 2003) and when considering emotional aspects (Boehner, 2005).

There are, of course, ethical issues associated with this kind of wirelessly-networked, ubiquitous technology (Stone, 2003), not least surrounding privacy. These are discussed in Section 2.3.2.

Ark and Selker (Ark, 1999) in their introduction to an IBM Systems Journal focused on pervasive computing, note that there are a broad range of areas for pervasive computing research. These include, but are not limited to: education, communications, infrastructures, input devices and social uses.

2.1.2 Computing and non-work environments

Domestic and care environments are very different settings to the workplace (Cheverst, 2003). Care must be taken when conducting research in these new areas to avoid blindly following the assumptions and methodologies associated with workplace-based research: for example, examining personal routines and environments can be seen as inappropriate and intrusive. That said, in some respects carrying out research can be easier in these environments. For example, buy-in is generally gained from all participants, not just the owners of a particular facility who require their employees to cooperate.

Some aspects of domestic and care environments are emotional. Work environments are generally regarded as just that: professional places where work gets done. These other environments, however, have very different uses, some of which are ‘worklike’ (the administration of balancing household accounts, for example, and paying bills and cleaning), and others which are completely different (such as: family meals, parties, television watching, game playing, and so on). In these environments, the primary aim is living, not productivity (Cheverst, 2001).

One piece of research looking at living and not productivity is by Howard et al (Howard, 2004), who investigated the support of intimacy between family members. They note the ambiguity of intimate communications, which convey emotions and feelings, and are very private. Electronic ‘gifting’ is described: this is the giving and receiving of messages of love and appreciation.

Technologies in domestic environments need not be staid or predictable: they can be used in creative and ludic ways. For example, Gaver et al (Gaver 2004) present the Drift Table, an electronic coffee table which displays slowly moving aerial photography. The movement of this photography is controlled by the distribution of weight on the table’s surface. Gaver et al used the table to investigate ludic activities: activities motivated by curiosity, exploration and reflection.

Another project of Gaver’s was the History Tablecloth (Gaver, 2006): this was a flexible screen printed with electroluminescent material to form a grid of lace-like elements. When objects were left on the table, cells beneath them lit to form a halo that grew over time, showing the flow of objects.

Anderson et al (Anderson, 2003) discuss technology to support ‘feeling close’ in affectionate relationships when people who are geographically distant. They discuss conveying non-verbal messages and augmenting objects with simple behaviours or digital traces that would hold their presence. For example, they discuss the giving of ‘beans’, which contain a secret object. Similarly, Keller et al discuss a system to increase the family communications via ambient awareness rather than spoken or written words (Keller, 2004).

Other work describes the provision of public information displays which are *enjoyable* and have art-like properties as well conveying information. (Ljungblad, 2003)

Paulos (Paulos, 2003) discusses the importance of play, illustrating its relevance to pervasive tasks such as blogging, tagging and message play. As can be seen, re-imagined functionality need not be limited to standard use of standard technologies.

'Smart home' projects may also be considered. One such project is the Aware Home (Kidd, 1999), which consisted of two identical living spaces. Aims included investigation of context awareness and ubiquitous sensing, and individual interactions with the house. Another project is the Millennium Home (Dowdall, 2001), aimed at elderly users who are not cognitively or physically impaired, but are at risk of becoming ill or injured through the course of home life.

Other work has considered technical, social and pragmatic challenges involved in providing such technologies in the domestic environment (Edwards, 2001). Edwards and Grinter list seven such challenges, including issues such as setting up and maintaining technologies, and reliability.

Understanding user requirements in care settings (such as care homes and hospices) can pose methodological challenges (Cheverst, 2003). It is necessary that designers know not only what they are designing and what it should do, but also who will use it. Designing for people just like the designers themselves may exclude a host of people, including the disabled and those in care.

Cheverst et al (Cheverst, 2003) carried out research in a hostel for former psychiatric patients. They used several techniques, including ethnographic study, user-centred design and evaluation, and cultural probes (Gaver, 1999). These methods helped Cheverst et al to build two systems: one provided medication reminders, and the second allowed patients to send an alert when in danger or distress (Cheverst, 2001). The cultural probes, consisting of informal, friendly packs with items such as maps and postcards, help overcome some of the 'professional distance' between researchers and users. Whilst designing a system to support inhabitants taking responsibility for their daily medication, various ethical aspects arose, which are discussed in Section 2.3.2.

Dewsbury (Dewsbury, 2001) has considered the social and psychological aspects of smart home technology specifically within the care sector, noting that assistive technology does not replace personal care. Recommendations include taking a long-term view of the inhabitant's condition, considering emotional aspects, and considering all stakeholders.

Daniel et al (Daniel, 2009) present a discussion of using computing assisted technology in the home environments of elderly persons. When presenting this work, they raised the issue that professionals can perceive the need for additional technological devices while the elderly would not see this need.

The majority of existing work on integrating technology with everyday life focuses on the office environment (Rodden, 2004). The home is a very different environment to that of the office, with a focus away from 'tasks' and 'procedures' and more towards 'routines'.

2.1.3 Home routines and design

Understanding use of technology in the home facilitates its design. Technology at home is often intertwined with routines. For example, Crabtree and Rodden (Rodden, 2004) often encountered situations where a household member would watch a specific program on the television (or listen on the radio) before leaving for work, or carrying out some other daily activity.

Crabtree and Rodden examined 22 family homes, and comment that the process of carrying out research in the home is little different than carrying out such studies in the workplace: the logistics involved and issues such as the Hawthorne effect (Mayo, 1933) are similar. They found that there are many information resources in the home involved with producing and consuming communications. One example which they study in depth is the handling of postal deliveries: who picks these up, where they are placed, who opens them, and so on.

A similar piece of work by Hughes et al (Hughes, 1998) considers a series of studies of domestic environments, and the design challenges they raise. The information was gathered over the course of six-month long ethnographic studies at various homes in the UK. It was noted that the placement of technology within the home reflected the daily routines of inhabitants. For example, the majority of households had mobile or portable TVs and stereos, allowing the distribution of their functionality at will. Participants were often wary of new technologies, citing high running costs, a lack of time to use them, or envisioning their children over-using them. Disruption was also seen as a dissuading factor.

Spaces are designed to support particular activities, usually multiple activities. A room may perform as a study, playroom, guest room or games room, depending on its configuration. It was noted that aesthetics are important in all of this. Hughes et al conclude that binding a technology to a space can limit what use can be made of the space; a commitment which should not be undertaken lightly.

Rodden and Benford (Rodden, 2003) have considered the form of buildings, and the implications of this for pervasive technologies. They note that one can understand the domestic setting through ethnographic studies, longitudinal studies and design-based methods. Devices can take the form of information appliances (standalone interactive devices with specific functionality), interactive household objects (where interaction is incorporated into the object's form – for example, picture frames as displays), or augmented furniture. The first of these three forms (information appliances) is the most intrusive.

Rodden and Benford draw upon Stewart Brand's (Brand, 1994) model of buildings, in which buildings are composed of six areas:

- Site: geographical setting and legally-defined boundaries
- Structure: the foundations and buildings
- Skin: exterior surfaces
- Services: utilities built into the building – wiring and plumbing
- Space plan: the interior layout – positioning of walls, ceilings and doors
- Stuff: furniture and possessions

(Note that Brand's model is American. The equivalent UK model counts site, structure and skin as one item, 'shell', because UK legislation tends to view these three things as one.)

It is noted that research so far has not focused on the site, structure and skin of buildings, instead focusing on stuff, space plan and services. That said, the former three dominate the latter three, influencing their placement and operation. Work involving digital artefacts in the home is in the stuff and space plan areas, while that involving environments and infrastructure operates in the services area.

People are constantly altering their home environments, and the different 'layers' of a building interact. For example, the location of services (power and aerial points) affect the placement of television sets. Issues arising include supporting services, cueing the presence of pervasive systems (which lack physical wires), and predicting how inhabitants will engage with the services and their maintainers.

2.2 Social Technologies

It is necessary to understand existing social technologies in order to consider how to broaden access to these. Various academic works consider social aspects of technology: for example, Foth (Foth, 2006) discusses how internet technologies can impact social formations of urban residents, while Ridings and Gefen (Ridings, 2004) explore why people join virtual communities.

Some work focuses on elderly users. Donaldson (Donaldson, 2005) describes two devices to facilitate companionship and discussion between co-located elderly people: these are the 'TeleTable', used to arrange and organise digital media, and the 'Pitara', used to associate physical mementos with digital media. Keyani (Keyani, 2005) presents a dancing environment to provide elders with exercise, entertainment and social engagement.

Other technologies focus on linking families. Mynatt (Mynatt, 2001) present the 'digital family portrait', a device to convey an impression of a family member's daily life over the last month, accounting for health, relationships, activity and events. Plaisant et al (Plaisant, 2006) present shared family calendars, a method for multiple generations of a family to share their calendar information as a tool to increase connectedness.

Similarly, Sellen et al (Sellen, 2006) describe the whereabouts clock, a device to support awareness of people's location and activities in an office environment. Initial evaluation suggested that it helped people have a virtual presence, locate others, and feel they belonged in a group of colleagues.

The rest of this section focuses on social networking websites, first describing what is meant by the term, and then describing their history.

2.2.1 Social networking websites

There are varying definitions of social networking websites: boyd and Ellison (boyd, 2007) define them as sites which allow users to construct a public or semi-public profile; articulate a list of users with whom they share a connection; and view and traverse these lists. Meanwhile, Golbeck et al (Golbeck, 2005) suggest that dating sites such as match.com are not social networking websites, while other sources (Abram, 2008) suggest they are. boyd and Ellison (boyd, 2007) also note that social sites are often about expressing existing connections, rather than meeting strangers, and are primarily organised around people rather than interests: this is a contrast to Usenet and forums, which are traditionally organised around topics.

Social websites may have different foci, for example towards blogging, careers, religion or general social interactions. For the purposes of this work, 'social networking websites' are sites which offer a specific focus upon augmenting relationships.

Although work exists examining the *functionality* of social networking sites (Richter, 2008), understanding the experiences associated with them remains a challenge. User experience of these websites appears to vary wildly, according to the specifics of the site and the way in which it is used. Types of use include searching for new friends, maintaining existing relationships and investigating people met offline (Hart, 2008). Sas et al (Sas, 2009) discuss the emotional experience of using the Facebook[2] website. They suggest that the main uses of Facebook are cooperative, and that the most memorable experiences are private communications between close friends and engagement in public performance while experiencing entertainment.

It would appear that existing, widely-accepted usability guidelines (such as Nielson's ten guidelines (Nielson, 1994)) are not applicable in this new area. According to those guidelines, Facebook is terribly designed (for example, its design is not "aesthetic and minimalist"), and yet its success suggests that it is certainly not hard to use (Hart, 2008). Understanding the modern-day web experience may help drive more holistic design guidelines.

Reaching this understanding may not be simple. Experiences which superficially appear similar may manifest differently in different media. This is demonstrated by Dix's use of deconstruction to translate the experience of pulling Christmas crackers from a physical to a digital domain (Dix, 2003). Deconstructive approaches offer a method for understanding user experience, and are explored further

[2] http://www.facebook.com/, last accessed December 2010.

throughout much of the rest of this thesis. Chapter 3 presents further material on user experiences.

2.2.2 The history of social networking websites

boyd and Ellison (boyd, 2007) suggest that the first social network site was SixDegrees.com, which launched in 1997: the site allowed users to build profiles, list friends and (from 1998) browse friends lists. These items of functionality were available on different sites before 1998, but not in combination. SixDegrees closed in 2000, while from 1997 to 2001 various community tools became social sites by adding support for profiles and publicly listed friends. Figure 2-1 shows boyd and Ellison's history of launch dates for many major social sites (and re-launch dates of community sites with social features), according to their definition of such sites (see Section 2.2.1).

Since 2006, the take-up of social networking sites has continued to increase. Twitter[3], a lightweight microblogging platform, launched in March 2006. By January 2009, the BBC reported[4] (based on material from industry analysts HitWise) that UK usage of Twitter had rocketed by almost 1,000% over the past year. Meanwhile, a list of major active social sites on Wikipedia[5] included 197 items, as of April 2009. This non-exhaustive listing excluded niche social sites, such as those made possible by services such as Ning, a website which enables users to build their own social sites.

It is difficult to discern the dates at which specific items of functionality (e.g. photo-sharing, or public and private messaging) appeared on specific sites. It is possible, however, to consider the origins of social networking sites. Not all such sites begin as such: for example, QQ began as a Chinese instant messaging service, LunarStorm as a community site, and Skyrock as a French blogging service: each of these later added social networking features (boyd, 2007). Similarly, various ethnic community sites (AsianAvenue, MiGente and BlackPlanet) re-launched in 2005 - 2006 with social networking features and structure.

---

[3] http://twitter.com/, last accessed December 2010.

[4] http://news.bbc.co.uk/1/hi/7844595.stm, accessed November 2010

[5] http://en.wikipedia.org/wiki/List_of_social_networking_websites, accessed November 2010

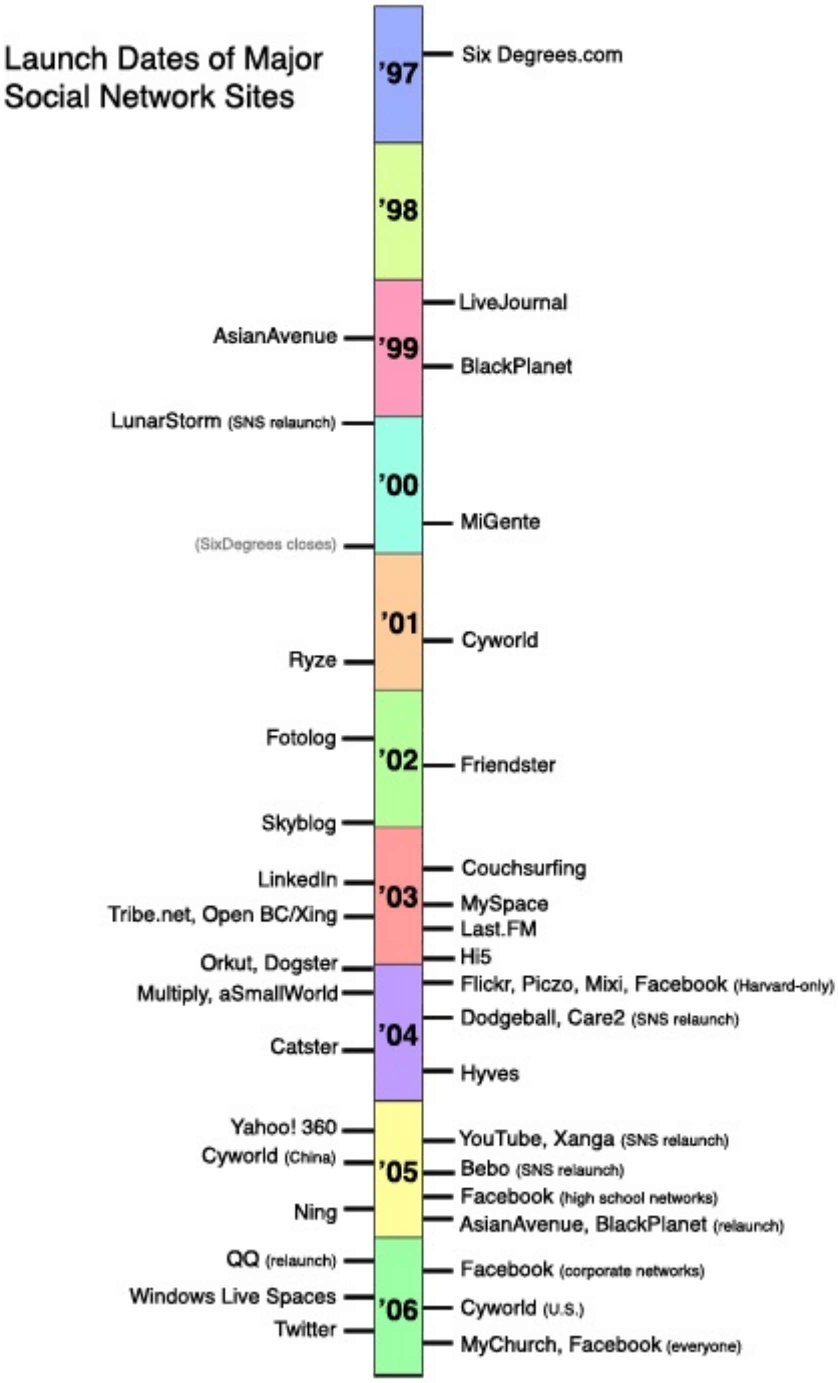


Figure 2-1 Timeline: launch dates of major social sites (boyd, 2007)

As can be seen in Figure 2-1, many new social sites launched after 2003. Most of these were profile-centric, but professional sites such as LinkedIn also appeared, as did interest-specific sites such as Dogster and MyChurch. Finally, as social media and user-generated content bloomed, media sharing websites added social features - these include Flickr, Last.FM and YouTube.

In recent times, mobile versions of social sites have become increasingly available: for example, in late 2010 Facebook launched various location-based services such as ‘Deals’, which enables users to receive deals from nearby participating merchants[6]. This functionality is inherently different to that of the web-based social experiences of Facebook, where geographic location does not affect the stream of data from friends. This difference is highlighted by Humphreys (Humphreys, 2007), who discusses Dodgeball, a location-based social networking site: he remarks “to refer to Dodgeball strictly as a social network site may be misleading”, and goes on to discuss how Dodgeball offers a limited web component yet the majority of access to the service is via a mobile phone.

The implication of this is that although mobile social services exist, they are not equivalent to web-based social services, suggesting that the problem of providing equivalent social experiences to those available on the web remains unsolved.

2.3 The Impact of Ageing, and Design for the Elderly

The suitability of a computer system’s interface is always important: even the most useful and efficient system is worthless if people cannot use it correctly. The appropriateness of a system’s interface is particularly important when the people who will use it are not able-bodied – for example, the elderly, small children, or those with health problems (Newell, 2004).

Accessibility is also of particular importance when dealing with a pervasive system, due to the novel nature of that interface: for example, one could ask whether there is a standard, textbook interface for a medical temperature-monitoring system or consider the difficulties of interaction with smart homes (Dowdall, 2001). Some work has been conducted in this area: one example is presented by Brewster (Brewster, 2002), who was concerned by the usability of small screens. He suggested the use of sonically-enhanced buttons to augment visual cues: tests in the usability laboratory found that this worked very well, although testing in more realistic situations such as walking outside found the improvements were not quite so strong.

[6] http://www.pcworld.com/printable/article/id,209675/printable.html, accessed November 2010

This section discusses user-focused design, describes the impact of ageing and issues in design for the elderly, and examines ethical issues in this area.

2.3.1 Ageing

The UK population in 2004 was 59.9 million, with a median age of 38.6 years (Royal Society, 2006). It has been broadly predicted that the population will reach 65 million by 2021, with a median age of around 41 or 42 years. The proportion of the population aged over 65 years is expected to have risen from 16% to around 18% by 2021. Although elderly people are by no means the only group with generally poorer access to technologies, they are certainly a significant one (Namazi, 2003). Given this significance, it is important to consider the effects of ageing, and the relevance of these effects to this research.

A widely acknowledged psychological change that comes with age is the decline in cognitive processes, especially memory (Mather, 2005). Not all types of cognitive process decline with age: for example semantic memory (memory of meanings and understandings) typically increases or remains steady. As such, recall of words studied a few minutes previously has been shown to decline over a four-year period, but implicit memory of recently studied words does not show a decline with age.

Shock (Shock, 1951) carried out a great deal of work looking at ageing. He notes that hearing, vision and motor responses are all affected by ageing: loss of hearing is greater in males, and the loss tends to affect the higher tones more. Farsightedness tends to occur more with ageing, and the elderly require a greater minimum level of light. Motor responses are slower: a very gradual decline begins between 40 and 50 years of age. There is a large range within this: the performance of the fastest third of the eldest group in a study equalled the average performance of the young group.

Shock also noted that average scores on intelligence tests diminish with increasing age: vocabulary-type scores are good, but those involving numerical computation, series completion, picture arrangement and so forth all featured a significant decline. The most difficult areas were those where subjects had to break away from old mental habits and adapt to unfamiliar situations. Shock notes, however, that the elderly have slower responses and more knowledge and experience, all of which biases the results of such tests.

Jacko examined the effects of multimodal feedback on the performance of older adults with different visual abilities (Jacko, 2003). Results suggests that non-visual (auditory or haptic), multimodal feedback resulted in performance gains over visual feedback, both for users with normal sight and those with impaired sight.

The elderly cannot be simply classified as one group: they are hugely diverse, particularly given the extensive and varied life histories and experiences that define them. Nonetheless, certain physiological changes, described above, tend to be found in us all as we age. These must be considered when designing computer systems for the elderly.

Uptake of technologies by the elderly must be considered. Gill (Gill, 1985) notes that the elderly are generally more resistant to change than younger members of society, and are thus more cautious about adopting new technologies. Namazi et al (Namazi, 2003) delve deeper, noting a wide-ranging set of obstacles to this uptake. Obstacles may be physical, cognitive, personal, technological, organisational or environmental. It is important to ensure that the elderly are not pressured to take on new systems in which they have no interest.

Several theories from Information Systems research could be relevant. The first is the Technology Acceptance Model (Davis, 1989), in which perceived usefulness and ease of use determine an individual's intention to use a system. The second is the Unified Theory of Acceptance and Use of Technology (Venkatesh, 2003), in which usage intention and behaviour is influenced by user expectations of performance and effort, as well as social influence and facilitating conditions. This model is the unification of eight prior models, including the Technology Acceptance Model.

Various research considers the design of computer systems for elderly people:

Perry et al (Perry, 2004) note that only 10% of older people live in supported accommodation and that characteristics of the elderly population tend to include forgetfulness, hearing and motor limitations, and a fear of crime. Consistency is important in systems for this user group, as different interaction methods can be confusing. User preferences are likely to be broad, due to differing backgrounds.

Pfeil (Pfeil, 2007) notes the increasing use of virtual communities online and the importance of making these technologies accessible to people, such as the elderly, who have different needs. Wiley (Wiley, 2006) presents the design of a simple, pen-and-paper based interface to allow elders access to email. Hirsch (Hirsch, 2000) describes a four-month study into the experience of elders and caregivers, and concludes that social, emotional and environmental factors all play a key role in the eldercare experience and the adoption and use of new products: this is another argument for a holistic approach to design.

Newell and Gregor (Newell, 2004) discuss specialist and mainstream design for older and disabled people. They note that older users are more likely to have one or more disabilities, as well as different wants and needs than able-bodied users. Those with permanent disabilities are only a fraction of the population of people with reduced functionality: everyone can be temporarily disabled by accidents, alcohol, stress, fatigue or their environment.

There is a difference between traditional user-centred design for able-bodied users, and that for user groups including those with disabilities. Newell and Gregor have proposed 'user-sensitive inclusive design', where older and disabled users are included as informants on research and development teams. They also note the value of narrative methods such as story-telling metaphors, and attention to the aesthetics of design.

2.3.2 Ethical implications

It is important to be aware of the ethical implications of any research. These aspects are particularly relevant when dealing with elderly or vulnerable users, using technology of a potentially invisible nature, and accessing highly sensitive information relating to health or security around the home.

The ethical implications of pervasive technologies attract attention due to fears over issues such as data theft. Stone (Stone, 2003) discusses some issues. One example used is the EZ Pass, which allows users to automatically pay highway tolls, speeding up traffic: yet use of such a pass allows the tracking of cars as they pass through toll gates. Wearable tracking devices for children are discussed: if a child is kidnapped or lost, these devices would be extremely useful. As Stone points out, however, it is unlikely that a teenager about to go out on a date would hold such a viewpoint. Others are yet more concerned about these technologies, with a Professor of applied ethics saying "We are building an infrastructure for totalitarian control."

Stone (Stone, 2003) cites another Professor who suggests a future where people embed devices which could, for example, increase their elasticity, letting them jump higher: he says this will threaten our sense of what it means to be human. He does not discuss existing augmentations, such as glasses, pacemakers and false teeth.

A major issue when providing assistance via technology is appropriately empowering users and supporting independent living, rather than simply creating a dependence on the technology (Cheverst, 2001) (Cheverst, 2003). This is exemplified by work facilitating the daily living routine and skills of former psychiatric patients in a hostel (Cheverst, 2001): it was necessary to build a system which would reassure patients (who worried about forgetting medication), but leave the task of managing medication in their own hands, rather than doing it for them. Awareness of the individual care needs and social implications of the technology is deeply important, as is awareness that technology cannot simply perform tasks for people without potentially creating dependence. The intrusiveness of the technology is another issue to consider in this context (Dowdall, 2001), especially in situations where the users have a lack of computer experience.

Ting (Ting, 1999) notes that the issue of confidentiality is growing in prominence as use of computing and communication technologies spread. These concerns are not new: Weiser (Weiser, 1989) noted over two decades ago that one rogue device could record everything in a room. Various legislation, such as privacy and freedom-of-information acts, has been introduced, although it has been said that US legal safeguards on privacy are inadequate, fragmented and inconsistent (Gostin, 1997).

Introna and Pouloudi (Introna, 1999) note the need to disclose information for the benefit of some people, and to safeguard the privacy of other people by not doing so. They describe this tension as being between privacy and transparency. They note that privacy is hard to define, with no universally accepted definition existing. Definitions range from 'the right to be left alone' (Brandeis, 1890) to 'control over knowledge about oneself' (Fried, 1968).

Another important aspect is intrusiveness, which is a noted issue (Ramchurn, 2004): providing the right information on the right device at the right time and with the right level of intrusiveness has been discussed previously (de Roure, 2005).

2.3.3 Summary of ethical issues

The ethical issues which have arisen highlight considerations that are difficult to support in the traditional design and engineering process. These include:

1. **Resistance to change and wariness of technology**

The elderly are generally relatively cautious about adopting new technologies (Gill, 1985), and can be intimidated by computer systems (Namazi, 2003). Namazi et al note that computers are not designed for operation by frail individuals and elderly persons with physical or mild cognitive impairments. As such, will elderly people want to use new technologies? How tolerant will they be of design problems in such systems?

Use of existing technologies with which the elderly are familiar (for example, televisions and phones) can provide a possible way forward. It is important that systems are not intimidating, with a minimal learning curve: Weiser (Weiser, 1989) wrote about being able to pick up a computer and use it as you would a piece of paper. Technology acceptance models such as those described in Section 2.3 could help predict the outcomes when new users are presented with technologies.

2. **Emotional impact and perceptions**

It is important to consider the emotional impact of any technologies introduced. Technologies which superficially increase the social connectedness of an elderly home-owner, for example, may in fact increase their isolation: relatives who have assumed responsibility for checking they are well may feel a reduced need to ‘check in’ with them. It is important to examine people’s perceptions of the systems with which they interact.

3. **Studies outside the workplace**

There seem to be few studies looking into people’s interactions and use of technology outside the workplace. In general, they appear to be avoided, and seen as inappropriate and intrusive, despite Cheverst et al’s (Cheverst, 2003) successful use of ethnographic study, user-centred design and cultural probes to elicit useful information in these environments. It would appear that such studies need not be inappropriate or intrusive: indeed, in some respects they can be easier than workplace-based studies. For example, as a rule the entire household agrees to take part, in contrast to a workplace study where individual employees may not be consulted beforehand.

4. Intrusiveness

The question of intrusiveness is highly relevant: for example, monitoring someone's location raises many issues. Would people agree to this monitoring? How would it affect their daily lives? Who would be able to access the data?

2.4 An Example Social Infrastructure

This section presents an example system built by the author in response to the motivation described in Section 1.1 and grounded in the theory presented earlier in this chapter. This section describes the motivation behind the work and prior work on information flow and messaging. The prototype system is then presented, followed by a scenario-based evaluation of the system. This section closes with a summary of the system.

2.4.1 Motivation

To attain the goal of greater availability of social technologies it is necessary to decouple information from its original modality. For example, the content of an email is the text, which as well as being displayed on a computer monitor could be printed out, displayed on a television screen or vocalised via a telephone with text-to-speech technologies.

Decoupling content from modality allows greater flexibility in terms of *what* information can be received *when* and *where*. An implementation would enable a user, Alice, to stream her voicemail to her PC if she had forgotten to bring her mobile phone to work, and mean that she could email her grandfather Derek, even though he may not own a PC. Derek could access direct messages such as emails and SMSes, and also more ambient social information (e.g. microblogs or shared photographs) from his family members. Figure 2-2 (the current situation) and Figure 2-3 (the vision) show how a multimodal infrastructure can enable this decoupling.

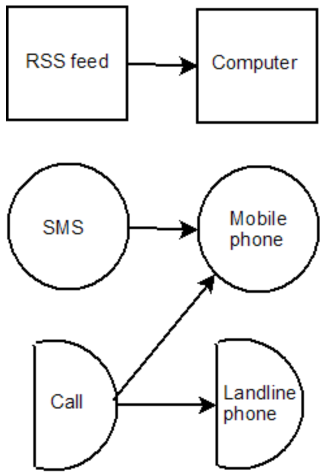


Figure 2-2 Message modality constrains which devices can receive messages

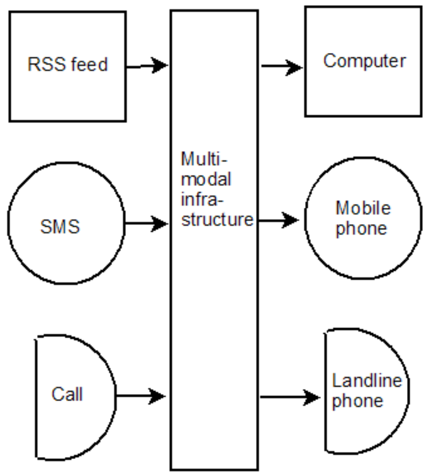


Figure 2-3 Content is decoupled from modality and can be sent to any device

To provide this functionality in an appropriate way, any infrastructure able to choose the most appropriate modality must be able to account for a person's context – for example, their location, current activity and priorities (which may concern mode of communication and intrusiveness).

Several components are needed to realise this vision. Firstly, an underlying pervasive messaging model allows transport of information between locations, and translation between modalities. This system must understand channels such as email, IM and audio data. Secondly, a formal user information model (user proxy) describes a user such as Alice, her relationships with others, preferences for modality, and rules (such as not being phoned between 11pm and 7am).

The envisioned system can use any suitable technologies to determine user location: this might include the current cell of a mobile phone, the Wi-Fi network used by a PDA, or GPS data. Physical location must be mapped to something meaningful in the context of the user, for example 'work', 'home' or 'park'. In the envisioned system, users may also explicitly notify the system about their current location. In the developed prototype, locations and subscriptions are initially set up in an XML file, and may be edited via the GUI.

2.4.2 Literature: information flow and messaging

Various proposed and implemented systems have looked at pervasive communication infrastructures. For example, the Mobile People Architecture

(Maniatis, 1999) embodies a vision where people, rather than their disparate devices, are the endpoints of communications. A 'personal proxy' tracks user locations, and accepts, converts and forwards communications as appropriate. The concept of Universal Communication Systems, which combine various modalities of communication (Andrews, 2001), is not a new one. Examples include: email notifications about voicemail (Liscano, 1997); 'consoles' for group conversations via SMS, email, IM and the web (Heyer, 2008); and a proposed system to route emails and phone calls dynamically according to user context (Kamioka, 2004). Similarly, Nakanishi et al have prototyped a system to redirect calls and emails based on people's schedule, location and available devices (Nakanishi, 2002).

The Iceberg architecture (Wang, 2000) aims to integrate cellular telephony networks and the internet. The Universal Inbox (Raman, 2000) uses this to redirect communication based on pre-defined user preferences. Active Messenger (Marti, 2001) routes email to pagers, phones and faxes, based on calendar and other contextual information. It allows users to define preferences according to their location and the time of day. Another implementation, Mercury (Lei, 2004), integrates phones, IM, email and pagers. Prior work by IBM includes LAMA (Location Aware Messaging for Accessibility)[7]. LAMA conveys location-relevant information such as train announcements or aeroplane boarding calls via smart phones.

Despite this work, practical deployments have been slow, perhaps due to the challenges of implementing this technology in the wild. Turk (Turk, 2005) notes the need to integrate channels and address privacy issues, whilst Branco (Branco, 2001) raises questions such as what data helps ascertain user context, and how best to map content for impaired users. Other issues include negotiating cooperation across mobile networks and the cost of setup (in terms of time) for users.

Recent developments in social communication, such as microblogging and instant photo-sharing, have introduced new requirements to multimodal messaging systems. We are still in the process of understanding user attitudes and behaviour on popular platforms such as Twitter, and social networking websites such as Facebook or MySpace[8]. Unlike email, SMS and IM, these technologies are not primarily about

[7] http://mqtt.org/projects/lama, last accessed December 2010.

[8] http://www.myspace.com/, last accessed December 2010.

direct messages, but a more ambient awareness and broadcast, and so they must be treated differently in the context of multimodal communications.

The messaging prototype described here differs from existing designs and implementations in several respects. Firstly, the vision isn't only about routing direct communication, but also information about the wider world (RSS feeds[9] and sensor data, e.g. electricity usage or car mileage) and personal data (reminders and ambient awareness of friends' activities and wellbeing, achieved with text and photographs through social networking mechanisms such as Facebook).

Thus, the system can route two types of data:

1. Personal data, access to which requires authentication (e.g. direct communications such as email and text messages and ambient communication such as social networking data)
2. Publicly available regularly-updated material, such as Twitter streams, blog posts, sensor data and other items on RSS feeds

The system allows different levels of notification, from viewing new information only upon explicitly logging into the system, to being woken in the night when an urgent message arrives. The system has a very broad audience: users may own PCs or mobile phones, but they need not. Anyone with a device which can interact with the system would be able to use it, including owners of older technologies such as televisions or landline telephones. Additionally, the emphasis on the assistive nature of this technology naturally leads to an approach of inclusive design, involving stakeholders where possible (Newell, 2004).

2.4.3 Prototype

The current prototype (Owens Hooper, 2009b) uses IBM's Lotus ® Expeditor micro broker[10] to convey messages and is coupled with a simulation environment. The messaging system is fully functional, while the simulation environment consists of a GUI that describes the movement of messages in the underlying system and allows

[9] Really Simple Syndication: a web feed format used to allow subscribers to pull regularly updated information (such as blog posts or 'feeds' of photographs).

[10] A small message broker for integrating parts such as applications, sensors, and actuators. It is designed to be deployed on small, appliance-type devices, often at locations remote from data centres. See: http://tinyurl.com/36vewar, last accessed December 2010.

users to send new messages. The system is not connected with message endpoints such as RSS feeds or LCD displays. The prototype demonstrates the soundness of the underlying logic and model, and enables exploration of scenarios.

The system receives incoming information (which in the future will be from websites, email inboxes, sensors etc) and delivers it to an appropriate endpoint. End users shouldn't have to concern themselves with the type of a message: whether the content was sent as an SMS, email or Tweet need not affect when and how it is received. Currently people may choose a different modality for message sending according to message priority or knowledge of the recipient's context. For example, if Alice's friend Bob is in the cinema, she may send him a text message but not phone him. This system aims to eventually automate that rules-based logic.

A screenshot of the current simulator can be seen in Figure 2-4. The simulator provides a listing of information from the world as modelled by the system: this list describes the time, devices, people, locations, data sources and events within the world. The right hand side contains controls for changing the state of various items within the world: it is possible to change a person's location and subscriptions, to send messages, and to change the time in the world. Below these controls are a change log (which records button presses, which result in state changes in the simulator) and a message log (which records the result of publishing messages).

These controls allow users to see how messages traverse the simulated world. By changing the subscriptions and location of a person, the user changes which messages that person will receive, and where. For example, as shown in the screenshot, it is possible to walk through scenarios, such as viewing what happens when Matt publishes photos on the stream to which Gerald is subscribed. Figure 2-4 shows the simulator loaded with the default XML model of the world (which includes the Matt and Gerald characters): the simulator GUI provides mechanisms to move the characters, change their subscriptions, and send messages.

Note that aspects such as the current time and events can affect message receipt. When the system doesn't know where a person is, it checks that person's calendar to see whether they are currently attending an event: if so, it reasons they are at the event's location, and tries to send the message accordingly. If no suitable devices are available upon attempted message delivery, the proxy of a given person will queue the message for later delivery.

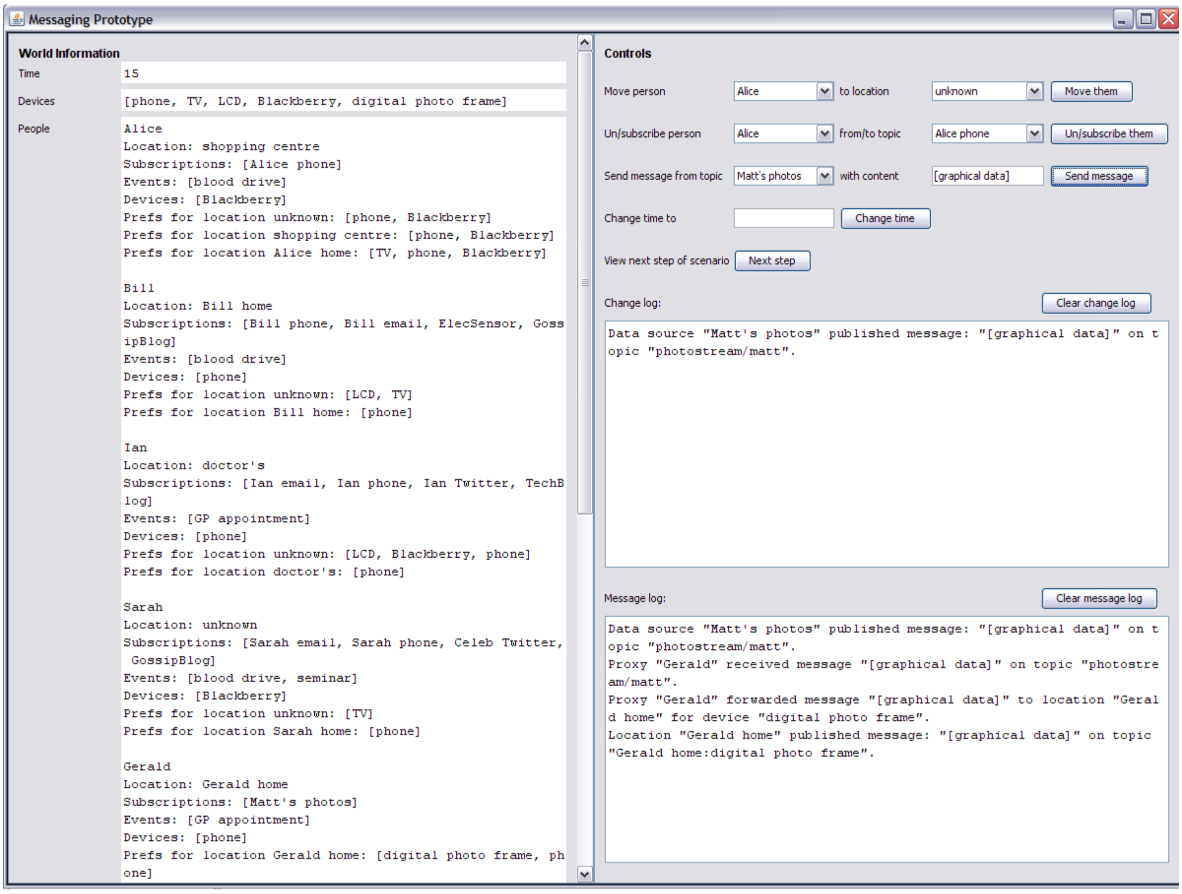
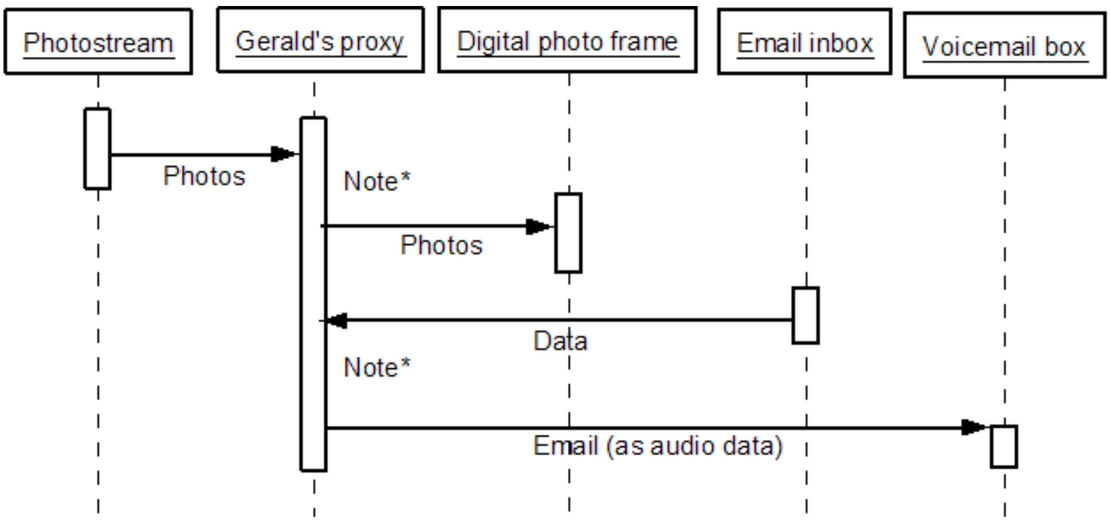


Figure 2-4: Screenshot of simulator

Figure 2-4 showed the simulator demonstrating the beginning of the given scenario, in which Gerald's proxy transforms and routes incoming material based on his preferences.



*Note: At each of these points, the proxy checks Gerald's preferences, and reasons that, based on Gerald's current location and preferences, it is appropriate to send the data on to the sources shown.

Figure 2-5 Scenario sequence diagram

2.4.4 Scenario-based evaluation of prototype

This section demonstrates that the underlying multimodal model is fit to support the scenario described in Section 1.1; describes the scenario built into the prototype; and discusses how the prototype can be extended to support more sophisticated scenarios.

Section 1.1 provides an initial scenario to demonstrate the original vision. The prototype system supports the functionality described in this scenario, which includes:

- viewing data from an online photograph stream on a digital photo frame
- receiving email content via a telephone or other device (note: the device type 'microprinter' is not incorporated into the current model, but adding this is trivial. See later in this section for more on this)
- sending voice messages without needing to know how or when the recipient will access these
- following online updates via a television display

A scenario was included with the prototype's simulator, designed to be used with the '`worldinfo.xml`' file. This file describes a world including the character Bill, a few locations he may travel to, his subscriptions, preferences and devices. Its purpose is to easily demonstrate how the system might affect one person in the course of a day.

The scenario is:

> *It is 8am, and Bill is at home. An email is sent to Bill, which is published as text to his mobile phone (his preferred – indeed, only – listed device for receiving messages at home). 9am comes, and Bill is now at his workplace, IBM. The morning passes, and midday arrives. At this point, a phone message arrives for Bill, which is routed to the LCD screen at IBM: the LCD is chosen as Bill hasn't listed preferences specifically for IBM, but the LCD is top of his default list.*
>
> *More time passes. At 1pm, Bill's location changes to unknown: he has left IBM, in fact to head to the university, which is holding a seminar at 2pm. Come 2pm, Bill's home electricity sensor publishes an update. Although Bill's location is still set to unknown, the system reasons that he is at university, as it knows from his diary that he plans to attend the seminar between 2pm and 3pm: the sensor update is delivered to the LCD at the university.*

Now it is 4pm, and someone sends Bill an email. Unfortunately, the system has no idea where Bill is: his location is set to 'unknown', and he is not attending any current events. The message is queued for later. When the time gets to 5pm, the system again tries to send Bill's message, checking his events to see if his location can now be inferred: it cannot, and so the message remains in Bill's queue. Note that Bill has a phone on his person throughout this: however, according to his preference lists, he only wishes to receive phone messages when at home.

Now Bill goes home, and the waiting email message is transmitted to his phone. At 7pm the 'GossipBlog' data stream publishes an update. This is propagated to Bill's phone (and incidentally, also to the Blackberry of Sarah, another character who is also subscribed to this stream).

The scenario can also be viewed step-by-step in the simulator, where each new step is triggered with another click of the 'next step' button:

1. Set time to 8am
2. Set Bill's location to Bill home
3. Publish data on "Bill email" stream
4. Set time to 9am
5. Set Bill's location to IBM
6. Set time to 12pm
7. Publish data on "Bill phone" stream
8. Set time to 1pm
9. Set Bill's location to unknown
10. Set time to 2pm
11. Publish data on "ElecSensor" stream
12. Set time to 4pm
13. Publish data on "Bill email" stream
14. Set time to 5pm
15. Set Bill's location to Bill home
16. Set time to 7pm
17. Publish data on "GossipBlog" stream

Note that pressing the 'next step' button again after the final step will trigger a message in the change log to the effect that the scenario is ended, and pressing the button again will restart the scenario.

The prototype can be extended to support more sophisticated scenarios. Consider the following:

Bill leaves his desk at work for a meeting as his house broadcasts some data about electricity usage. Bill has locked his desktop computer, so his proxy knows he is not present. It consults his laptop and calendar, and reasons that he is on his way to a meeting in room B12, due to start in five minutes. Bill has marked sensor data from the house as non-private, so the system flashes his name on an LCD screen embedded in the wall outside room B12. Bill soon passes the screen and sees his name; he is slightly early for his meeting, so he stops to press a button and view the data. A passer-by notices, but this is unimportant as the information is not private.

During the meeting, Bill's wife Janet updates her status on Facebook. This is not a high priority item, but as Bill's laptop is open, the system flashes up an alert about the update. As no one else is looking at Bill's laptop, he reads this personal information.

Soon after, Bill's boss Abi emails him. Bill's proxy reasons that this may be relevant to the current meeting, but his laptop is now closed. The proxy sends a high-priority message as an SMS to his phone, which vibrates. Bill reads the SMS, which displays the subject line of Abi's email: it is not about the current meeting, so he ignores this for now.

The current prototype cannot fully support this functionality. Six changes are required to enable the scenario outlined above, listed below.

1. Proxies may access computers associated with an end user, in order to find out whether the user is active at a computer
2. Proxies need awareness of messages' sensitivity and whether output devices are public (e.g. a communal LCD in a corridor is public; a mobile phone is not)
3. Allow a person's subscriptions and messages to be marked as 'high priority'
4. Proxies are aware of the relevance of events to non-attendees (e.g. reasoning that Bill's boss is connected with his current meeting)
5. Proxies reason that just before or after an event, an attendee is likely to be in transit to or from that event
6. Allow output devices to notify a subscriber of message receipt, only displaying the message content when prompted by the recipient

Appendix A: Code Changes to Improve the Social Infrastructure) details how and where the functionality for each change would be implemented. In summary however, two types of modification are required to incorporate the above functionality:

1. The `Proxy.sendMessage` method, the reasoning engine of the system, needs to handle the more complex logic
2. Classes which represent objects in the world need to be augmented. For example, the `Message` class requires a Boolean about priority and the `Location` class needs a Vector listing adjacent locations. Some new classes are needed, for example to represent `Devices` and `Subscriptions`: these are currently represented as simple Strings. A microprinter device, mentioned at the beginning of this section, would also need to be added to the device representations.

None of the required changes impact the framework of the system: the changes affect only the complexity of existing logic (in the `sendMessage` method) and objects within the model.

Enhancements towards the envisioned system are generally straightforward. Including detail about time, people and events in preference lists and enabling 'trusted contacts' involves increasing the sophistication of the `PrefList` and `Proxy` classes; similarly, enabling preferences for environments and events involves no fundamental change to the framework.

Other, further-reaching extensions might include enabling preferences based on time as well as location, and transforming data into an appropriate format for the end modality. Future work might include an exploration of how to deal with conflicting user preferences and issues of privacy and intrusiveness.

From the above, it can be concluded that the existing prototype provides a suitable and extensible architecture for the envisioned messaging system, and currently uses a simple world model and reasoning. It can be built upon to create a pervasive infrastructure with a social layer, allowing the easy integration of pervasive social tools and modalities.

2.4.5 System summary

The prototype messaging system represents a framework to route and transform messages. It was built as an infrastructure to support a social fabric: an interface and social model. Clean separation of the infrastructure and the social fabric allows appropriate focus on each part: the underlying messaging infrastructure can easily be used in other domains, while the social fabric can be developed as a discrete concern, with due consideration given to issues such as how it will be used, the interface and human-computer interaction aspects, and ethical considerations.

The design presented here will allow provision of fluid multimodal information to its users, based on context, priorities and preferences. This is motivated by the ideal of allowing people without use of digital technologies to access (and return) communication and social information which originated with these.

This proposed system is capable of delivering information in appropriate formats according to various factors such as personal preference, location and which channels are available. The system releases people from the burden of choosing a communication channel based on what they imagine is best for recipients, such as worrying over whether to send an SMS to someone who may be asleep, or whether to call someone who could be at the cinema.

2.5 Conclusions

This chapter has introduced pervasive technologies, which involve the availability of many 'invisible' devices distributed through the environment. Issues include privacy and security, while the novel nature of these technologies mean that usability issues are especially important, due to both their novelty and their diverse group of users. Computing in non-work environments, particularly homes and care homes, has been described. Playfulness can be an important aspect of these: existing ludic systems include digital family portraits (Mynatt, 2001) and table-based devices such as the TeleTable (Donaldson, 2005) and the Drift Table (Gaver 2004). However, there appears to be a dearth of creative design methods which are particularly suited to work with this kind of technology.

This chapter also described design for the elderly: the different needs and desires of elderly or disabled users should be considered, whilst keeping in mind their vulnerability and different (very broad) profiles. When considering ethical aspects, issues include those associated with the use of pervasive technologies for vulnerable users. Other issues of interest include the wariness of technology often exhibited by elderly populations; emotional impact and perceptions; studies outside the workplace; and intrusiveness of systems. The area of accessibility is rich, and work has considered accessibility for elders and accessibility in the context of pervasive technologies.

Existing socially-oriented work considers virtual communities, devices to facilitate face-to-face social interactions and devices to increase the sense of presence for geographically disparate work or family groups. Social networking websites come in varying forms and appear to support a rich array of online interactions: however, the provision of social networking experiences off the web appears to be little-addressed.

In summary, the literature suggests the importance of suitable interfaces, user-focused design and a holistic approach which accounts for non-technological aspects including users' comfort and social interactions. Pervasive technologies offer a way to provide rich experiences in an accessible fashion; the provision of social aspects would be a good use of this space.

As observed in Section 2.4.2, there exists prior work looking at multimodal communications, much of which is from at least several years ago. Recent developments in social communication, such as microblogging and photo-sharing, have introduced new requirements to these systems, and we are still in the process of understanding user attitudes and behaviour on popular platforms such as Twitter, and social networking websites, such as Facebook or MySpace. Unlike email, SMS and IM, these technologies are not primarily about direct messages, but more ambient awareness, and so they must be treated differently in the context of multimodal communications.

The next logical step from this work is to design the interface of the envisioned system described in Section 2.4. The underlying infrastructure is clearly able to convey materials, but it is clear that the manner in which they are presented is critical to the experience of end users. Although the functionality of the proposed system is clear, the delivery mechanism is not, due to a lack of understanding of user experience.

The problem with current approaches to redesigning experiences for new contexts (such as moving from the web to pervasive computing) is that there is no framework for understanding what it is about these experiences that people find valuable. Without that understanding, creating a system to accurately capture those benefits becomes very difficult.

In order to move on, it is necessary to consider what motivates people to make use of social networking websites, what it is they offer, and how that experience might be transferred to a pervasive environment. A purely technology-driven approach would result in a technologically functional system, without evidence that it will re-produce the social experiences which we seek to support.

Chapter 3 presents existing work in the area of understanding user experiences and supporting creative software design.

Chapter 3 Creative Design and User Experience

Chapter 2 showed that it is necessary to understand user experiences in order to build systems to effectively address problems such as those set out in Chapter 1: providing alternative modes of access to experiences that are currently web-based is problematic. Deeply understanding those experiences offers an approach to the issue.

To that end, this chapter explores three areas. Firstly, because redesigning experiences is a creative process and we are interested in the field of Computer Science, Section 3.1 presents a discussion of creativity in the context of software design. Secondly, the field of User Experience (UX) is explored in Section 3.2: its recent emergence is discussed, and models of experience and methods for understanding experience are presented. This chapter also discusses techniques for evaluating creative design methods in Section 3.3.

3.1 Creativity and Software Design

This section introduces the concept of creativity and its relevance to software design, before discussing software design itself, along with the lack of traditional tools for understanding experiences.

3.1.1 Creativity

There are many competing definitions of creativity, ranging from 'a pure mystery' (Gero, 2002) to 'an innate talent that can be trained' (Candy, 1999). Silva

(Silva, 2009) presents an excellent overview of this area, providing definitions which range from creativity as a mysterious and personal phenomenon to something which is rationally interpreted and social. A simple dictionary definition is 'the ability to produce new and original ideas' (Summers, 1987).

There exists a strong research community in the area of understanding creativity and design, evidenced by the existence of conferences such as Create[11], ACM Creativity and Cognition[12], and Creativity and Innovation in Design[13]. (Shneiderman, 2007) remarks that "The growth of interest in creativity support tools in recent years is gratifying" and notes the emergence of various workshops and funding.

Understanding the abstract creative process helps us to interpret methods which scaffold creativity: there are various models of creativity. Wallas presents a model consisting of preparation, incubation, illumination and verification (Wallas, 1926), while Osborn presents a two-phase process of idea generation and idea evaluation (Osborn, 1963). More recently, Shneiderman focused on the importance of disseminating creative output, contributing a model in which people collect, relate, create and donate materials (Shneiderman, 2000). Shneiderman also presents three perspectives on creativity: inspirationalist (in which preparation and incubation lead to moments of illumination), structuralist (emphasising more orderly approaches and the study of previous work) and situationalist (focusing on social and intellectual context).

What is common to the models is the inclusion of two phases:

1. A divergent stage in unconstrained brainstorming (or other techniques) are used to generate ideas.
2. A convergent phase in which the proposed solutions are critically evaluated.

(Shneiderman's model includes these two phases within the 'create' step.) This concept is illustrated in Figure 3-1.

Examples of specific techniques are brainstorming (Osborn, 1963), a group creativity technique designed to generate multiple ideas in response to a problem; TRIZ (Straker, 2002), a problem-solving tool with roots in a survey of patent literature; and the Creative Problem Solving Process (CPS) (Reali, 2010), a technique

[11] http://www.create-conference.org/, last accessed December 2010.

[12] http://dilab.gatech.edu/ccc/index.html, last accessed December 2010.

[13]http://www.desire11.id.tue.nl/, last accessed December 2010.

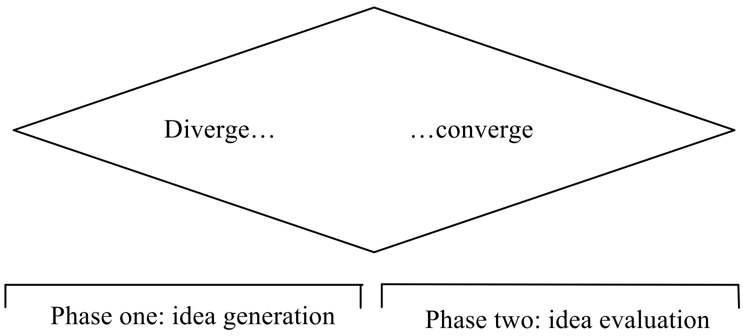


Figure 3-1 Model of creativity

developed by Osborn and a colleague. Creative techniques support the creative process by providing specific steps. For example, CPS involves three overall phases:

1. Understanding the challenge
    a. Find objectives (identify the goal)
    b. Find facts (gather relevant data)
    c. Find problems (clarify which problems must be solved to reach the goal)
2. Idea generation (generate ideas to solve the problem)
3. Idea evaluation
    a. Find solutions (move from ideas to implementation solutions)
    b. Find acceptance (plan for action)

Titchener (Titchener, 2005) notes the creative nature of Computer Science, saying "The field has exhibited immense creativity, ranging from innovative hardware such as the early mainframes to software breakthroughs such as programming languages and the Internet… computers and the theory of computation have provided great opportunities for creative work."

As described, there are general techniques for solving problems in creative ways. There are also creative techniques specific to the fields of Computer Science and Software Engineering, which the author has previously explored (Hooper, 2009, Hooper, 2010c). Examples of such methods include architectural patterns (Gamma, 1995), storyboards and sketch prototypes (Curtis, 1990) and scenarios and personas (Cooper, 2007). Many other examples exist[14]. Additionally, Software Engineering methods can be used in creative, agile ways. For example, the UML (Unified

[14] A list of methods in the field of user-centred design is presented here: http://www.usabilitybok.org/methods, last accessed December 2010.

Modelling Language) is often applied creatively by Computer Scientists in ways which are appropriate to the problem at hand.

Table 3-1 presents some existing literature in the field of Software Engineering and creativity. The examples include work in the areas of requirements elicitation, software design and evaluation, and how software is used in the field.

Area	**Examples**
Introduce creative methods	(Grudin, 2002) describes personas, fictional users representative of groups of real-life users of systems. They can be used in participatory design and to consider social and political aspects of design.
	(Gaver, 1999) introduces cultural probes: informal, friendly packs with items such as maps and postcards. These are used to elicit a broad set of data on end users.
Discuss new ways to use creative methods	(Panke, 2007) describes the use of personas as a tool for evaluating e-learning web-portals by incorporating the method into a model of quality assurance.
	(Tedjasaputra, 2004) presents pair writing of scenarios, in which scenarios are written by two writers working in unison, for example a usability specialist and a software developer. This usage enables pairs to explore and reach a common understanding of requirements and design ideas.
Understanding creativity	(Warr, 2005) presents a review of definitions of creativity and proposes a unified definition, and also compares social creativity with individual creativity.
	(Terry, 2002) considers how people use software tools in creative, experimental ways, and discusses how modern user interfaces do not support such practices but impose a linear progression which is not suited to creative pursuits.
Supporting creativity	(Hewett, 2005) examines how to design computer environments to support creativity and innovation, focusing on key factors to consider in the development of such environments.
	(Maiden, 2004) discusses how to encourage the creative aspects of requirements engineering, which traditionally involve elicitation, analysis and management but little creation or invention. They describe workshops held to encourage creative thinking in this context.
	(Ross, 2010) proposes abductive reasoning to support creativity, noting that brainstorming efforts are not always successful and suggesting that increased support for fostering creativity may help.

Table 3-1 Existing work on creativity in Software Engineering

3.1.2 Software design

Many definitions of design exist (Atwood, 2002), with varying foci. For example, definitions may hinge on the initiation of change, the nature of the problem, or relations with materials or people (Silva, 2009). One simple definition is that design is a process that produces a new or refined product (Warr, 2007).

Much research has considered software design: for example, 25 years have passed since Gould and Lewis (Gould, 1985) discussed principles of system design, and it is 15 years since design patterns were first appropriated by Computer Scientists (Gamma, 1995). More recently, Sengers (Sengers, 2005) suggested the use of critical 'reflective design', in which designers analyse how technologies reflect and perpetrate unconscious cultural assumptions. She also describes participatory design, value-sensitive design, critical design, ludic design, critical technical practice and reflection-in-action.

Gould and Lewis (Gould, 1985) describe three principles of system design which allow production of a useful and easy-to-use computer system. These are:

1. Early and continual focus on users.
2. Empirical measurement of usage.
3. Iterative design whereby the system (simulated, prototype, and real) is modified, tested, modified again, tested again, etc.

Although these principles are straightforward, their evidence suggests that the principles are not always intuitive: 447 designers were asked to list five major steps in developing and evaluating a new system. Only 16% mentioned all three items, and 26% mentioned none.

Design has been discussed in the context of specific domains including the home (Hughes, 2000), eldercare (Hirsch, 2000), hypertext (Nanard, 1995) and multi-modal, ubiquitous systems (Perry, 2004). Topics include metaphors (Madsen, 1994) and aspects such as aesthetics and narrative (Newell, 2004).

Although understanding user experiences is valuable, traditional methods do not always enable this. For example, tools such as scenarios and personas (Cooper, 2007) focus on end users, but it is not always clear how they scaffold an understanding of user *experiences* grounded in specific contexts. Hart (Hart, 2008) considers the contrast in results of traditional heuristic evaluation and a study intended to elicit user experience: according to the traditional evaluation, Facebook is terribly designed (for example, it is not "aesthetic and minimalist"), and yet its success suggests that it is certainly not hard to use. Understanding the modern day web experience may help drive more holistic design guidelines.

It is also possible to 'borrow' design methods from other disciplines. Actor-Network Theory (ANT) is one method which could be turned to software design. It emerged from the work of Callon (Callon, 1986) and Latour (Latour, 1987), and models the flow of interactions and processes between ‘actors’, which may be people or artefacts. ANT gives a process-based perspective on interactions between people and concepts, and insight into how networks grow and are used to achieve goals. It can offer different perspectives, for example modelling a bank as a network or as an entity within a larger network.

However, existing work suggests that ‘borrowing’ and blindly applying theories from other fields is inappropriate, and that it is necessary to understand them deeply in order to apply them suitably (Johnston, 2001). This does not preclude Computer Scientists from working with Sociologists or other domain experts to apply such theories, but it is a word of warning for appropriating theories without due consideration.

The emerging field of User Experience offers an alternative way to consider experience while designing software systems.

3.2 User Experience

This section presents the field of User Experience (UX), first explaining the history of the field, before presenting models of experience and methods for understanding experience.

3.2.1 The field of user experience

The discipline of Human-Computer Interaction (HCI) has evolved over time, moving through three ‘waves’ or paradigms, first grounded in engineering, then oriented around process, and finally moving to user-orientation (Harrison, 2007). The first paradigm saw interaction as a man-machine coupling and was inspired by industrial engineering and economics. The second paradigm was influenced by cognition, drawing parallels between human information processing and computational signal processing. Finally, the third paradigm contains a set of perspectives and approaches with a central metaphor: all action, interaction, and knowledge is embodied in situated human actors. In other words, how we understand the world, ourselves, and interaction derives from our location in a physical and social world, highlighting subjective experience.

The third wave is exemplified by work concerning multiple meanings in design (Sengers, 2006) and approaches to 'reflective' design (Sengers, 2005), including areas such as participatory (Plaisant, 2006) and ludic design (Gaver 2004).

Sengers (Sengers, 2003), a proponent of the third wave, argues that although the analytic approach inherent in engineering provides one way to model human experience, it misses the rich complexity of human behaviour. She posits that the design of rich, meaningful experiences requires stepping away from a Taylorist task-based approach and taking a multi-disciplinary approach to experience.

Although there are many arguments for experience-oriented design, understanding experience is a demanding task. Goffman discusses the dynamism and multiplicity of human interactions (Goffman, 1959), comparing these interactions to a performance in a theatre, shaped by audience and environment. This is reflected in later work on social networking, which observes that people construct a social identity that they present, staging themselves for a particular audience (Richter, 2008).

Suchman also discusses the complexity of human interactions (Suchman, 1994), noting their reliance on context, assumptions and prior knowledge, while Nack (Nack, 2003) discusses the difficulties of capturing experiences in dynamic, interactive environments, again noting the necessity of understanding context. Bardzell (Bardzell, 2008) discusses the subjectivity of experience, and emphasises the need for rigour, to 'transcend anything-goes subjectivism and offer systematic, evidence-based analyses of subjective phenomena'. McCarthy and Wright discuss the importance of empathy, and the use of empathy to understand the relationship between user and designer (Wright, 2008), and van Harmelen (van Harmelen, 2001) remarks that "A complete set of descriptions can never fully describe the richness of the users' world."

Despite the challenges in this area, it is clear that the understanding of user experience has become an important topic. This is exemplified by, for example, the recent Special Issue of Interacting with Computers, entitled "Modelling user experience - An agenda for research and practice" (Law, 2010). Other indicators are the addition of 'understanding' to the HCI lifecycle (Harper, 2008), the occurrence of CHI workshops on Experience Centered Design (Blythe, 2006) and observations from the field about this new focus on incorporating multiple points of view (Sengers, 2006).

The growing field of UX reflects increased interest in user-focused approaches to Software Engineering (Hassenzahl, 2006). Traditional methods do not always enable the understanding of user experiences: for example, Hart (Hart, 2008) considers the contrast in information yielded by traditional heuristic evaluation and a study intended to elicit user experience.

Emotion is part of UX, and was raised by Picard (Picard, 1995), who argued for machines that relate to, arise from or deliberately influence emotion or other affective phenomena. Boehner (2005) and Hook (2008) are proponents of this view, seeing emotions as constructed in interaction. Others have discussed how to measure emotion: for example, Desmet suggests the use of both verbal and non-verbal methods (Desmet, 2003). In addition to describing methods to measure emotion, Desmet notes the relevance of unpleasant as well as pleasant emotions, pointing out the popularity of rollercoasters and frightening films.

Issues of accessibility are particularly relevant to this thesis. Previous work has discussed User-Centred design (UCD) in this context: Wiley (Wiley, 2006) applied UCD in designing a message centre for elders living at home, while Keyani (Keyani, 2005) applies the paradigm to design an augmented dancing environment for elders.

As discussed, understanding experiences is necessary when re-providing them in new domains. This problem is particularly relevant given the rise of Web Science, which emphasises the need for an interdisciplinary approach to holistically understanding the impact of the web: in the same way that many disciplines must work together to understand the web, re-providing affective, social experiences in different contexts requires the input of more fields than Computer Science.

3.2.2 Models of experience

Various frameworks of UX exist:

Forlizzi (Forlizzi, 2004) discusses user-, product- and interaction-centred approaches to understanding UX. She considers user-product interactions as falling into one of three categories:

1. Fluent: automatic and well-learned, e.g. riding a bicycle.
2. Cognitive: product-focused, e.g. learning how to use a foreign toilet.
3. Expressive: helping the user form a relationship with the product, e.g. setting a background image on a mobile phone.

She also identifies three types of context for experiences:

1. Experience: a stream of self-talk which might occur while walking.
2. An Experience: an event which can be articulated, for example going to see a film.
3. Co-Experience: creating meaning and emotion together through product use, e.g. interacting with others at a museum.

Forlizzi suggests multidisciplinary teams can use the framework to understand and generate interactions and experiences in their product and system designs.

Relating to the book 'Emotional Design' (Norman, 2003), Hassenzahl (Hassenzahl, 2003) notes the subjectivity of UX, and that experiences with products may vary between individuals, situations and over time. He describes two attributes of products:

1. Pragmatic attributes, concerning functional, utilitarian aspects.
2. Hedonic attributes, concerning aspects such as pleasure and interesting or exciting functionality, content or interactions.

Mahlke (Mahlke, 2005) considers hedonics and aesthetics alongside affect and emotion. He notes that traditional usability approaches focus on tasks and goals, efficiency and cognitive information processing. He identifies three types of non-instrumental quality, which are 'hedonics', 'aesthetics' and 'pleasure and fun'. He then presents four dimensions of experience:

1. Perceived usefulness
2. Ease of use
3. Hedonic quality
4. Visual attractiveness

Mahlke notes that to study UX of interactive systems it is important to take a holistic, integrative approach.

Wright and McCarthy (Wright, 2003) describe experiences as holistic, situated and constructed, and present a parallel framework which compares four relational elements of experience, and six aspects about how people make sense of experiences. The first part of the framework is the four strands of experience:

1. Compositional: what it's about, how it is composed, structure, consequences, explanations.
2. Sensual: look and feel, sensations like thrill, fear, excitement, welcomeness, belonging, unease.
3. Emotional: for example, anger, joy, disappointment, frustration. This is affected by feelings and not senses: "We can engender emotions associated with achievement through the exercise of control over sensations such as fear or anxiety […] Emotions are not just passive responses to a situation."
4. Socio-temporal: the time and place of an experience.

The second part of the framework concerns how we make sense of experience. McCarthy and Wright note that sense-making is reflexive and viewed through a person. The components of this part of the framework are:

1. Anticipating
2. Connecting (first impressions)
3. Interpreting
4. Reflecting (placing a value on the experience)
5. Appropriating (relating the experience to our sense of self, personal history and hopes for the future)
6. Recounting (finding new meaning by retelling and drawing out evaluations from others)

Wright and McCarthy note that experience is about what people bring to the situation as well as about the artefacts involved in that experience, and reason that we therefore can design for experience if we understand our users and are skilled and sensitive. They provide their framework as a way to see, discuss and analyse experience.

Later, Wright, Blythe and McCarthy characterise their approach with three themes (Wright, 2006):

1. A holistic approach to experience wherein the intellectual, sensual and emotional stand as equal partners in experience.
2. Continuous engagement and sense-making wherein the self is the centre of experience, is engaged in experience and brings a history of meanings and anticipated futures to complete the experience through sense-making.

3. A relational or dialogical approach wherein self, object and setting are constructed as multiple centres of value with multiple perspectives and where an action, utterance or thing is designed and produced but can never be finalised since the experience of it is always completed in dialogue with those other centres of value.

These models have varying, complementary foci: Forlizzi emphasises fluency and context, Hassenzal and Mahlke consider hedonics, and Wright and McCarthy focus on 'strands' of experience and sense-making.

3.2.3 Understanding experience

Traditional software design methods such as scenarios and personas (Cooper, 2007) can help focus designers on end users, but it is not always clear how strongly they scaffold an understanding of user *experiences* as grounded in specific contexts. There exist various methods to understanding UX, some of which have been previously explored by the author (Owens Hooper, 2009a). Methods may attempt to elicit opinions, experiences or emotions from users: see Table 3-2 and Table 3-3.

Method	**Description**
User studies (Kazmer, 2008)	Collecting qualitative semi-structured interview data about Internet-based research topics, for example via telephone, email, face-to-face and instant messaging.
Cultural probes (Gaver, 1999)	Send participants informal, friendly packs with items such as maps, postcards, cameras and booklets. Include very open-ended questions, such as 'what is your favourite device?' or 'what place does art have in your life?' Elicit attitudes to life, cultural environments and technology.
Videotaped Activity Scenario (Little, 2009)	Videotape scenarios developed from interviews, use as a prompt to generate detailed group discussion of experiences, beliefs and expectations.
Repertory Grids (Fallman, 2010b)	Elicit 'conceptual structures' (single dimensions such as warm-cold, work-leisure) and explore them and their inter-relations.
Experience Prototyping (Buchenau, 2000)	Encourage participants to engage in role-play to understand, explore and communicate ideas about experience.
Experience deconstruction (Dix, 2003)	'Deconstruct' a specific experience to elicit its properties, artefacts and effects.

Table 3-2 Approaches to eliciting user opinion and experience

Method	Description
Self-assessment manikin (Isomursu, 2007)	Pictures of puppets for measuring emotion along three dimensions (pleasure/displeasure, arousal, dominance-submissiveness).
Emocards and feedback applications (Isomursu, 2007)	Display cartoon faces to represent emotions: users select the one to most accurate reflect their current feelings.
3E (Isomursu, 2007)	Express emotions by drawing and writing.
Experience clip (Isomursu, 2007)	Users take short video clips in which they express their thoughts and feelings.

Table 3-3 Approaches to eliciting user emotions

Methods from other domains may also be considered. One example is phenomenography (Lewthwaite, 2008), a qualitative method used in education research to investigate how people experience or think about certain things. Branches of this field include narrative analysis, ethnography and the use of questionnaires. Another example is Actor-Network Theory, which was discussed in Section 3.1.2: as stated there, methods from external disciplines must be applied with great care.

Recent interest in experiential and emotional aspects has also led to work on how to evaluate systems in this new light. (Roto, 2008) gives one example of evaluation, while (Mahlke, 2007) presents a study on evaluation of non-instrumental qualities and emotions; Shen (Shen, 2007) describes intrusive and non-intrusive methods for evaluating information systems.

This section outlined methods for understanding experience. Traditional approaches to elicit user opinion (such as interviews and focus groups) have existed for a long time: the existence of the problem space which this thesis addresses suggests that these approaches (or their application) are not suited to the support of deep understanding of experiences for software design.

Newer methods such as cultural probes, role-play and video-based techniques and repertory grids exist alongside methods for eliciting emotions. The former approaches again aim to understand experience, but do not provide constructs for redesigning experiences in new contexts, and often do not include specific constructs for understanding abstract or emotional facets – or when they do (as in the methods for eliciting emotions), this is to the exclusion of other factors.

Experience deconstruction represents a technique which aims to address experiential effects and is paired with another approach (experience reconstruction) for redesign: however, this technique is presented in the form of a description of its application rather than as a step-by-step process. Like the other methods presented in this section, it does not include specific constructs aimed at understanding facets such as social or emotional elements.

3.3 Evaluation of Creative Design Methods

Methods of software design and experience-centred design both involve creative steps. The evaluation of subjective, creative methods is a non-trivial task. Shneiderman (Shneiderman, 2007) discusses the difficulty researchers face in striking a balance between qualitative and quantitative methods in this context. He observes that there is often pressure from journal and conference reviewers to provide statistically significant results, yet laboratory-controlled studies with many participants can be inappropriate. He draws attention to an increasing emphasis on close study of domain experts and the emergence of case study, observation and interview, intended to capture the processes that precede breakthroughs and understand how software design features promote creativity.

Fallman (Fallman, 2010a) presents a discussion of rigour and relevance in interaction design research. He defines ‘rigour’ as the validity and reliability of research: that is, whether it measures or explains what it sets out to measure or explain, and whether it is repeatable. He notes a difficulty with repeatability: it is highly unlikely in the field of creativity and design that different practitioners will create the same artefacts.

In terms of relevance, Fallman notes the weight of research contributions and generalisability, but primarily refers to Keen (Keen, 1991). Keen suggests that in order to be relevant, research must be:

- Interesting: addressing problems important to professionals
- Applicable: utilised by practitioners
- Current: addressing timely issues
- Accessible: understandable by practitioners

Fallman suggests that interaction design research can be conducted in different ways and for different purposes, suggesting three contexts: design practice, design studies and design exploration.

Kitchenham (Kitchenham, 1996) discusses methods of evaluating Software Engineering methods and tools. She lists nine types of evaluation: experiments, case studies and surveys (all of which may be qualitative or quantitative), qualitative screening, qualitative effects analysis, and benchmarking. She notes that evaluations can look at measureable effects of using a method or can establish a method's appropriateness.

Kitchenham notes that an issue with undertaking evaluations in industry is the possibility of sociological and managerial factors as well as technical difficulties: for example, if participants believe a new tool is superior, they are likely to gain good results; those results may not carry over to other software users who lack that confidence.

Furniss (Furniss, 2008) presents a study of how usability practitioners work in professional web design. The work includes a list of four factors of working practice which affect method usage:

1. Method fit with current working practice of practitioners and their organisations.
2. Fostering relationships with clients and colleagues.
3. Supporting communication of ideas.
4. Other aspects of working practice, such as reporting and visibility within communities. (It can take 5 - 10 years to have a full understanding of a method's fit in these contexts.)

Case studies offer another way to gain insight into design methods. Yin (Yin, 2008) describes case studies as empirical enquiries that investigate a phenomenon within its real-life context using multiple sources of evidence: case studies help to answer "how" and "why" questions.

A common concern about case studies is that they provide little basis for scientific generalisation. It has been noted (Shneiderman, 2007) that individual case studies can provoke multiple case studies in order to replicate findings with diverse users and problems; as multiple case studies replicate results, researchers gain confidence in the replicability and generalisability of cause-and-effect conjectures.

There exist various examples of case studies being used: Hertzum (Hertzum, 2003) applies them to use of scenarios, Minocha (Minocha, 2009) to experiences of social software and John (John, 1995) to use of the cognitive walkthrough method. John discusses the relevance of 'how' and 'why' questions in the field of HCI (referring to questions such as "How can a technique be used?" and "Why does a method work in this context?") although (Kitchenham, 1998) discusses the use of case studies to ask "Which is better?"

Shneiderman engages in a discussion of how to evaluate creativity support tools (Shneiderman, 2006): he notes that "Controlled experiments of specific features seem too narrow, as do gross comparisons of one tool versus another. Controlling for individual differences seems nearly impossible and specifying tasks is somehow at odds with the goals of supporting innovation or discovery."

He goes on to discuss the possibility of combining empirical studies with more naturalistic and creative situations, but seems to conclude that there are few examples of such work. He proposes the use of Multi-dimensional In-depth Long-term Case studies (MILCs) as one approach to this area, citing its multiple methods as a strength which enables the acquisition of multiple perspectives. He also notes the importance of enabling the achievement of users' goals within their work domain.

It would seem that there is a place for both quantitative and qualitative methods in this area, and that case studies offer a mechanism for understanding real-world usage of methods.

3.4 Summary

Software design, creativity and user experience are central to the subject of replicating technological experiences into new contexts. This chapter presented a discussion of these areas, introducing models of creativity and experience, and discussing software design and methods for understanding experience. It also discussed the evaluation of design methods, describing frameworks from Kitchenham and Furness, and the use of case studies.

As described in this section, there exist various methods (within HCI and from other disciplines) to understand user experience, varying from user studies to cultural probes to role-playing.

Existing methods for understanding experience involve eliciting information from end users to model their experiences. Excepting experience deconstruction, none of the methods surveyed integrated mechanisms for building designs based on the resultant knowledge, and although some methods focused on social and emotional aspects, no method appeared to consider both of these facets.

The evolving design space described in Chapter 2 is not well served by existing methods, which only sometimes consider accessibility and rarely account for physical or mental impairments alongside issues associated with pervasive technologies. When attempting to replicate experience, some aspects (such as social and emotional factors) are less well served by existing methods, if at all.

To meet the goal of broad access to social networking that was presented in Chapter 1, it is necessary to analyse the social networking experience towards translating the experience into new, accessible modalities. As discussed, Software Engineering lacks design methodologies oriented around understanding and replicating user experience: Chapter 4 presents a method developed by the author in response to these gaps.

Chapter 4 Creating TAPT

This chapter introduces the Teasing Apart, Piecing Together (TAPT) method, developed by the author for analysing experiences and redesigning them in new contexts. This was motivated by the work presented in Chapter 2, in which existing methods led to the creation of a messaging infrastructure to support pervasive equivalents of web-based social networking technologies, but did not provide support for understanding user experience towards translating experiential aspects into the new modality. It was also motivated by Chapter 3, which outlined existing work on User Experience and revealed a lack of methods to support understanding and replication of experiences.

The chapter opens with a description and analysis of 'experience deconstruction' (the original inspiration for TAPT). Section 4.2 describes the process by which TAPT was built, and the TAPT method itself. Section 4.3 presents an application of TAPT to aspects of social networking sites as an example of the method and its output, before closing with a discussion of the results of that study.

The work presented in this chapter has been published (Owens Hooper, 2009c) (Hooper Owens, 2009a).

4.1 Inspiration: 'Experience Deconstruction' as Presented by Dix

Alan Dix (Dix, 2003) first used deconstruction with Christmas crackers. Deconstruction is a design tool for understanding user experience, useful for providing equivalents to existing experiences in new contexts. The method is holistic and creative, rather than a traditional engineering approach.

A Christmas cracker is made up of an inner tube wrapped in brightly coloured paper. When pulled by two people, it splits into two uneven parts, making a bang as it does so (caused by a small exploding chemical mechanism called a cracker snap). Crackers generally contain a paper hat, a small plastic toy and a motto or joke.

Dix wanted to create virtual crackers on a website. Rather than trying to emulate real crackers, Dix succeeded in capturing aspects of the experience of pulling crackers, and translating those to the medium of the web. He did this by deconstructing the experience of pulling a cracker, and then reconstructing it in the new medium. By deconstruction, Dix refers to "taking apart, teasing out the strands that make something what it is … and, in this context, especially those that make something 'work' as an experience or as a designed artefact."

Deconstruction involves consideration of 'surface elements' and 'experienced effects'. Dix does not explicitly define these, but gives examples. An exemplar surface element of Christmas crackers is that they are traditionally 'cheap and cheerful': thus the webpage for virtual crackers was simple, with cheerful graphics. An aspect of the experience of pulling a Christmas cracker is the shared nature of the experience. To incorporate this, the virtual cracker system would not allow the sender to see the contents of the cracker until the recipient had 'pulled' it (by clicking on a link).

Dix's reconstruction involved a simple webpage with a web toy and joke, and a mask to print and cut out; virtual crackers are sent by email, and the sender cannot view the content until the recipient has. The page changes slowly, and plays a sound file of the bang, when it works.

4.2 Formalising Deconstruction

The author took Dix's description of deconstruction and went about systematically defining it as a process. The following section describes the approach taken, which included building a step-by-step description of deconstruction based on Dix's paper, introducing a word limit for the initial experience description, dividing 'experienced effects' into two types, describing reconstruction and supporting abstraction of the original experience.

4.2.1 Change one: a step-by-step description

The first step was to produce a list to describe the deconstruction process, so as to be able to describe the overall process as well as consider each individual step. The list, below, is extrapolated from the example deconstructions given by Dix, who does not describe the process in detail. Note that 'surface elements' and 'experienced effects' are phrases coined by Dix.

1. Describe the chosen functionality and the experience of using it
2. List surface elements of the experience, such as the nature of the design (e.g. 'simple') and the physical parts of the design (e.g. 'diagonal line')
3. List experienced effects of the experience, such as 'sharing', 'openness' or 'excitement'
4. Consider how to translate the surface elements and experienced effects to the new modality

The lists generated in steps (2) and (3) describe the experience in an abstracted manner, away from the constraints of the original modality.

Presenting deconstruction as a step-by-step process allows us to more easily examine types of surface element and experienced effect. It can be seen that Dix appears to consider artefacts and properties to be surface elements, which are largely nouns and adjectives. Experienced effects tend to focus upon the physical, emotional and social effect upon participants: these descriptors tend to be abstract nouns, noun/verb pairs and perhaps adverbs.

The emphasis on, in particular, emotional and social effects is rooted in existing literature. For example Blythe (Blythe, 2006) refers to emotional and social designs in the context of experience-centred design while Wright et al (Wright, 2003) describe experience as four threads: compositional (addressed by taking a deconstructivist approach), sensual (relating to look and feel, addressed by surface elements), emotional, and spatio-temporal (the experience as situated in space and time: this aspect is considered in a later evaluation of the method, described in Chapter 5).

4.2.2 Change two: a word limit for the initial experience description

In order to better understand the process, the author applied the re-described deconstruction process to the experience of browsing a social networking website (chosen for its relevance to the original problem, as described in Section 1.1). It was found that browsing such sites is a very rich experience with many facets, and that deconstructing it was problematic. By contrast, deconstruction appeared to produce relevant results when applied to facets of the experience such as browsing microblog updates or uploading photographs.

This led to the first additional change to the process: limiting the initial experience description to 200 words, or just under half a page of text.

This change was made because deconstruction involves focusing on one specific experience: if an experience cannot be succinctly described in 200 words, the practitioner is probably trying to deconstruct something too broad. The solution is to break the experience in question into several parts.

4.2.3 Change three: two types of experienced effect

The process of applying the emerging deconstruction process to several facets of social networking websites left the author aware of the variety of types of experienced effect. Although Dix's original deconstruction does not divide experienced effects into subtypes, two types of effect became clear: literal and abstract.

Literal effects are concrete in nature, such as a loud noise or broadcast of information. Abstract effects are not concrete, and tend to concern emotional and social aspects. Examples are surprise, connectedness and cultural connotations. Literal effects can lead to abstract effects, but either subtype of effect may exist without a corresponding other half.

Dividing experienced effects into two groups like this has the advantage of scaffolding the process of considering the experiential aspects of an experience. Considering experienced effects in two separate steps makes each step smaller and therefore less difficult. Additionally, considering tangible, concrete effects is generally simpler than working out abstract effects, so tackling this part of the task first is a gentle start to the process.

4.2.4 Change four: more description of elements and effects

The formalised deconstruction process was now beginning to emerge, and is given below. In addition to the changes described above, the list includes more detail:

steps two and three include further descriptions and examples of the type of item which practitioners would expect to identify. This is to help practitioners understand each step more fully, and was included on the basis of informal discussions between the author and colleagues.

1. Briefly (in no more than 200 words) describe the chosen functionality and the experience of using it
2. List surface elements of the experience. These are generally nouns and adjectives relating to the design, such as 'bold diagonal line', 'simple text box' or 'complex arrangement of photos'
3. List experienced effects of the experience. These focus on the physical, emotional and intellectual effect upon participants, and tend to be abstract nouns, noun/verb pairs and perhaps adverbs. They are literal (e.g. 'a loud noise', 'broadcasting information') or abstract ('excitement', 'shared')

4.2.5 Change five: describe the reconstruction process

By applying reconstruction to various facets of social networking, the author was able to articulate a step-by-step process, thus:

1. **Brainstorm**, considering the new context of implementation. One might consider modality, technology and scale.
2. **Using these ideas, build an example reconstructed scenario**.
3. **Check the reconstruction**:
    a. Are all desired elements included? (Choosing to omit some key effects from the original experience is fine)
    b. Were any unintended key effects introduced?
    c. Refine the scenario, repeat these steps as needed.

Step one was included as a primer, with the prompt about 'modality, technology and scale' intended to help scaffold initial thinking. Step two of the process was intentionally written without great detail, as it is at this stage that the practitioner must make a creative leap in redesigning the initial experience in the new context. Step three was included to help designers ensure they had rebuilt the scenario as intended.

4.2.6 Approaches to reconstruction

After the above work, an informal conversation between the author and Dix led to the conclusion that there are two possible approaches to the creative leap in reconstruction: a holistic approach where the practitioner tries to reproduce the lists of effects in the new domain, and an incremental, transformative approach, where the practitioner takes each effect in turn. After this insight, the author attempted both approaches to reconstruction in different contexts, resulting in the following descriptions of each:

The *holistic approach* involves a functional matching of the experienced effects with the new, reconstructed experience. Instead of making incremental adjustments, the designer attempts to match the list of effects with the capabilities available in the domain of the desired, reconstructed experience.

The *transformative approach* is:

1. Take one of the experienced effects in the source domain (for example, 'openness')
2. Choose an experience in the end domain which matches this effect (for example, 'chatting in a café')
3. Consider whether the chosen experience matches each effect of the source domain (for example, 'one-to-many communication')
4. If it does match, compare it with the next source domain effect. If not, adjust the experience in the end domain to allow for this ('shouting in a café')
5. Repeat these steps until the reconstructed experience satisfies all of the surface elements and experienced effects of the original experience.

Note: at step one in the above sequence, if a concrete experienced effect is chosen, it may be necessary to consider a corresponding abstract effect. For example, 'a loud noise' is irrelevant if re-imagining Christmas crackers for a deaf community: however, the corresponding abstract effect of 'surprise' can still be facilitated via other means. This illustrates the importance of context in reconstruction: it is necessary to understand the effects of 'a loud noise' when pulling Christmas crackers.

As reconstruction involves a creative leap, the author decided not to impose either of the above approaches on practitioners, but instead allow them to choose the approach most suited to them.

4.2.7 Change six: support abstraction of the original experience

After applying the new process, the author realised the importance of abstracting the starting experience description. This led to two additions to the deconstruction process description:

1. Adding ‘identification of key effects’ to the deconstruction process: here, the practitioner examines the lists of literal and abstract experienced effects to see which items are essential to the experience in question.

This is to help the practitioner identify what aspects of the starting experience are essential and should be included in the subsequent reconstruction. The author added this step after finding it helpful for focusing on the important aspects of a deconstructed experience.

2. Prompting the practitioner to describe the abstracted experience to produce a ‘distilled experience’. This second description is written in experiential terms: that is, it is not connected with the context and technologies of the original experience, but instead with experiential aspects.

This helps the designer think of the deconstructed experience in abstract, neutral terms. As well as helping designers distance themselves from the modalities of the starting experiences, this neutrality of language was emphasised as a tool for boosting the accessibility of designs: for example, describing communication in a neutral way rather than talking about ‘showing pictures’ would leave a designer more likely to support communication in ways which might not be only visual in nature. The author added this step to support designers in creating designs less ‘led’ by the starting experience.

4.2.8 Naming the new process

This approach needs its own name, particularly because ‘deconstruction’ only refers to the first phase of the process, excluding the reconstruction phase, and also because ‘deconstruction’ is a word with different meanings in many domains: it thus has subtleties of meaning which may be misleading to practitioners. The approach further described in this chapter is named TAPT, which stands for Teasing Apart, Piecing Together. ‘Teasing Apart’ refers to the analytical, deconstructive phase of the process, while ‘Piecing Together’ refers to the reconstruction phase, where the analysed experience is rebuilt in a new context.

4.2.9 Teasing Apart, Piecing Together

This section describes TAPT:

The first phase of TAPT, Teasing Apart, helps people understand and 'distil' the nature of an experience. The output is a table showing how the experience breaks down. There are five steps in this stage:

1. **Briefly describe the chosen functionality and the experience of using it**. For example, if we were Teasing Apart photo-sharing on Facebook, we might write: "Facebook allows users to users to upload and caption photos, which can be commented upon by the photo's owner or other users. Viewers can 'tag' friends in photos, adding metadata which links images with people's profiles."
2. **List the 'surface elements' of the experience**. These are generally nouns and adjectives relating to the design. For example: a somewhat complex photo upload process; the option to annotate images with text; the option to 'tag' images, indicating who is shown; the ability to view photos
3. **List 'experienced effects'**, which focus on physical, social, intellectual and emotional effects, and tend to be abstract nouns, noun/verb pairs and adverbs. There are two types:
    a. Literal: concrete items such as 'broadcasting visual information' and 'sharing past experiences'.
    b. Abstract: for example 'consolidate online identity', 'openness about past experiences', 'anticipation of discussion', 'reminiscence', and 'uncertainty about responses and audiences'. This step is important.
4. **Identify effects which seem especially important, unique or key**. For example, 'broadcasting visual information', 'presence in the community', 'openness' and 'reminiscence'.
5. **Describe the abstracted experience in a neutral sentence**. For example, mention 'broadcasting' information rather than 'playing' it, as the latter implies an audio-visual modality. One might write of photo-sharing: 'A way to share and annotate imagery from the user's past; their audience can access and annotate that imagery'.

The list of elements and effects will vary in length according to the experience considered, as will how many effects are key.

Phase two is Piecing Together: this phase is a creative tool for generating ideas, where users re-imagine a Teased Apart experience in a new context. There are many ways to re-provide an experience, so there are no 'wrong' answers. Output is a description of the experience in its new context. The steps are:

1. **Brainstorm, particularly using key effects**, considering the new context of implementation. One might consider modality, technology and scale.
2. **Using these ideas, build an example reconstructed scenario**.
3. **Check the reconstruction**:
    a. Are all desired elements included? (Choosing to omit some key effects from the original experience is fine.)
    b. Were any unintended key effects introduced?
    c. Refine the scenario, repeat these steps as needed.

Here is an example of TAPT, applied to redesigning photo-sharing on Facebook and Orkut[15] in the context of the home:

1. *Description: Facebook and Orkut both allow users to upload and caption photos, which can be commented upon by the photo's owner or other users. The sites both allow users to 'tag' friends in photos, adding metadata which links images with people's profiles. Both Facebook and Orkut allow photo upload, captioning and tagging. Indeed, it would appear that surface elements and experience effects are nearly identical between the two sites, as both offer basically the same functionality.*
2. *Surface elements*
    - *a (generally complex) photo upload process*
    - *the option to annotate images with text*
    - *the option to 'tag' images (indicate which contacts are displayed)*
    - *the ability to view photos*
3. *Experienced effects*

Literal:

- *<u>broadcasting information</u> (underlining indicates key effects)*
- *<u>sharing past experiences</u>*

Abstract:

- *presence in the community – consolidate online identity by adding more data*
- *openness about past experiences*
- *anticipation of <u>discussion</u> about these experiences*
- *<u>reminiscence</u>*
- *uncertainty about responses and (depending on privacy settings) audience*

Abstracted description: A mechanism for broadcasting information on past experiences, to facilitate reminiscence and discussion.

Reconstruction: A novel tabletop to display a sequence of photographs chosen by the table's owner. It incorporates a touchscreen interface to allow the owner to lay down photos and annotations. Passers-by can annotate photographs with commentary.

The above example begins to demonstrate TAPT's helpfulness: there is a contextual change from 'uploading photographs to a website' to 'a semi-public tabletop display', but the abstract experience description bridges the gap.

Dix's approach to Christmas crackers involved deconstructing an experience in the real world and reconstructing it in a digital context, the web. The author, by contrast, wished to analyse a digital experience (using the web for communication and

[15] http://www.orkut.com/, last accessed December 2010.

ambient social awareness of friends' activities via social networking sites) and rebuild it in a physical context, by providing that information and interaction via the modalities previously described. Section 4.3 describes the first steps towards this goal.

4.3 Initial Application of TAPT

This section describes a systematic application of TAPT to aspects of social networking websites. Section 4.3.1 describes a survey of the functionality offered by a range of social networking sites and analysis of which functionality is key. In 4.3.2, TAPT is applied to those key aspects of functionality, before Section 4.3.3 summarises and discusses the results.

4.3.1 Survey of social networking functionality

Analysing the overall experience of using social networking websites, which offer a plethora of communication and awareness tools, is a problematic task. This is because use of each of these appears to produce an experience perhaps greater than the sum of its parts. Teasing Apart or deconstructing the browsing of a social website is almost equivalent to deconstructing the browsing of the web: both comprise many nodes and connections, points of functionality and diversion, and huge quantities of data. Given this, several key aspects of functionality common to social networking were chosen.

The first step was to survey the functionality offered by a range of social networking sites, and analyse which functionality is key across this range. Abram and Pearlman (Abram, 2008) offer a description of social networking websites, and in the course of this list the following sites: Facebook, MySpace, Friendster, Orkut, LinkedIn, Windows Live Spaces, Bebo, Meebo, Match, and QQ[16]. These sites have various differences: for example, MySpace focuses on music, LinkedIn is designed for career-related networking, and Match is a dating tool.

[16] These can be found at, respectively, http://www.facebook.com/, http://www.myspace.com/, http://www.friendster.com/, http://www.orkut.com/, http://www.linkedin.com/, http://spaces.live.com/, http://www.bebo.com/, http://www.meebo.com/, http://uk.match.com/ and http://www.qq.com/.

The author examined the homepage of each site. Each offers an array of features and functionality: it appears that items linked prominently on the home page (links visible without scrolling down) are primary features, and other functionality is peripheral. Examples of peripheral functionality are ‘to do’ lists on Bebo, birthday listings on Facebook, and classified adverts on MySpace. (Note: Meebo simply provides Instant Messaging functionality, while QQ is in Mandarin, and thus beyond the reach of the author: as such, neither of these sites was further examined.)

It was necessary to normalise the language used by these websites, so as to produce a common terminology: for example, many sites include within profile pages a ‘scrapbook’, ‘wall’ or ‘whiteboard’. This is a space for friends (and the profile-holder) to leave notes, which are akin to a publicly-visible version of directed messaging. This analysis refers to the concept as ‘public messaging’.

Functionality common across the sites was: a profile; a microblog; a friends list; public and private messages; photos; applications; discussion groups (or ‘forums’ or ‘communities’); and news feeds. There were some exceptions: LinkedIn, Match and Twitter do not supply all of this functionality. This is presumably because the functionality is not always relevant to LinkedIn and Match (photographs are rarely relevant in careers-related discussion, and Match users probably don’t want messages to be public), while Twitter is a very lightweight social networking service.

Appendix B: Social Sites’ Functionality fully lists the primary functionality of each site.

Commonly occurring items in profiles are: name, picture, age, gender, relationship status, location, free text (e.g. to list interests or quotes), contact information (e.g. email address, IM username, URL), listings of schools/workplaces, and a current ‘status’ (a microblog). Again, there are exceptions, notably LinkedIn (which does not include personal data such as age and gender) and Twitter, which provides a very lightweight profile (not including relationship status, for example). Appendix C: Social Sites’ Profile Data lists the profile fields for each site.

Having elicited the key functionality of social networking websites, the next step was to examine the surface elements and experienced effects of this functionality, in order to abstract it for transfer to new media. This process is described in Section 4.3.2.

4.3.2 *TAPT in action*

This section discusses analyses of four types of social functionality across three social networking sites: functionality that was Teased Apart is microblogging, public messaging, photo-sharing and discussion groups. The first two items are similar in some respects, as both involve sharing a small quantity of text in a fairly public way. The application of TAPT clearly brought out the differences between these. Photo-sharing and discussion groups were chosen as they provide a contrast to the other two types of interaction.

The three social sites studied were Facebook, Orkut and Twitter. Facebook and Orkut offer fairly similar social functionality, and it was hoped that using TAPT would help show differences between the sites. Twitter is a lightweight social networking mechanism, selected as it would, therefore, provide a contrast.

The analysis is based on data collected in August 2008. Although now in 2010 photo-sharing is common on Twitter, at the time of data collection photo-sharing services such as Twitpic[17] were only beginning to emerge, and were not in mainstream use. As such, there is no analysis of photo-sharing on Twitter.

Each application of TAPT includes the analysis of the design and experiential aspects of the interaction (Teasing Apart), followed by a sample rebuilding of that experience (Piecing Together). This analysis was conducted after the first five steps of creating TAPT, but before step six, which involved indicating which effects are key and writing an abstract description: consequentially, the TAPT analyses do not include these two steps.

Analyses of public messaging on Facebook and Twitter follow, while all ten analyses can be found in Appendix D: TAPT Applied to Social Networking.

Public Messaging on Facebook (the 'Wall')

As described, social sites often provide a 'wall', 'whiteboard' or 'scrapbook', where friends and the profile-holder can leave notes. Variations include the ability to augment these notes with HTML formatting or images.

---

[17] http://twitpic.com/

Surface elements

- a box for up to 1000 characters of plaintext[18]
- a 'share' button
- a list of previous messages and activities by the person whose wall this is, most recent first

Literal experienced effects

- quick and easy
- communicating (one to one)
- being overheard

Abstract experienced effects

- social connectedness (conversing, or letting someone know they are in your thoughts)
- anticipation of a response
- uncertainty (will there be a reply? When? Who else will read the message, how will they respond?).

Piecing Together

A rebuilding of this functionality must take into consideration the above elements and effects. An implementation resulting from this analysis might provide a very simple, clean interface for entering public messages and clarity that the message is public (through showing previous messages left by other people, and perhaps on first use a brief explanation of the mechanism).

One way to implement this in a new context might be to install a microphone and speakers on the door of someone's office. Passers-by may press a button to record a message for the office's owner ("Hi Andy! I dropped by to chat about work, but you weren't here. Catch you later!"); the last ten messages are played in a repeating loop. A time limit (30 seconds, perhaps) reflects the character limits in the original medium. This mechanism is audible to people in the office area, just as public messaging on Facebook is visible to connections in the social network.

(Note: the above demonstrates a holistic approach to reconstruction, as described in Section 4.2.6.)

[18] Note: now, in late 2010, Facebook enables the posting of links, images and videos on the 'Wall'. However, at the time of this analysis, it did not.

Public Messaging on Twitter ('@replies')

Unlike Facebook and Orkut, public messaging is carried out from the home page of Twitter, where users view the stream of tweets from users to whom they are subscribed. By contrast, the other two sites offer public messaging on the profile page of the recipient of the message:

Public messaging on Twitter is achieved via use of the @reply mechanism: users type their message, which like a microblog entry is limited to 140 characters, but include the text "@username" to direct the message to the person with that username. For example, "@bill Are we still on for tea at 10?"

Surface elements

- a plaintext box for up to 140 characters of free text
- one 'update' button
- a list of previous messages, most recent first
- username of the recipient, and the @reply mechanism

Literal and abstract experienced effects

- quick and somewhat easy (additional load: users must understand the @reply mechanism and know the username of the recipient, although they need not navigate to the recipient's profile page in order to send a public message)
- communicating (one to one)
- being overheard

Abstract experienced effects

- social connectedness (conversing, or letting someone know they are in your thoughts)
- anticipation of a response
- uncertainty (will there be a reply? When? Who else will read the message, how will they respond?).

Piecing Together

When considering rebuilding, one must remember that like Facebook and unlike Orkut, public messaging on Twitter involves plain text only: additionally, updates are limited to 140 characters. Obvious changes to the existing reconstruction (the voice recorder on the office door) would therefore be to provide functionality to leave audio-only messages limited to, say, 15 seconds in length. This does not quite capture public messaging on Twitter, however, as one would still have to walk to the office door of the recipient, whereas Twitter enables messages to be left via its homepage.

Instead, the audio equipment could be in the reception of our office building, and people leaving messages must speak the name of their intended recipient for clarity.

(The above adjustment of the location of audio equipment demonstrates the incremental approach to reconstruction described in Section 4.2.6.)

4.3.3 Summary and discussion of results

Section 4.3.2 presented a systematic application of TAPT across a set of social networking sites and functionalities. This section discusses the results, which are summarised in Table 4-1.

It can be seen that the elements and effects associated with public messaging are very similar across all three platforms, although message richness varies according to the presence of HTML formatting and graphics, and message length. These differences are reflected in the various Piecings Together.

Analysis of microblogs and photo-sharing also yielded very similar elements and effects across the sites, and the Facebook and Orkut approach to discussion groups was similar. The Twitter implementation of discussion groups differed, being based on tagged microblog posts rather than users explicitly signing up to a particular community. In a similar vein, the Facebook and Orkut implementation of functionality for discussion groups was generally rather similar, while the Twitter approach differed, using a microblogging-based mechanism to achieve community.

It can be seen that Facebook and Orkut generally presented very similar experiences: this similarly can also be seen in Appendix B: Social Sites' Functionality and Appendix C: Social Sites' Profile Data, which show both sites providing similar functionality and profile breakdowns. The two sites do show some differences in their implementations of functionality, for example with Orkut's implementation of public messaging allowing richer material to be expressed.

One might notice that there is a blurred distinction between Orkut's public messaging (called 'the scrapbook' on the site) and Facebook's microblogs entries (or 'statuses'): the former allows 1024 characters of HTML-rich text with images, while the latter allows 420 characters of plaintext with URLs, images, videos and application-specific data. Unlike public messages on Orkut, a Facebook microblog entry is constrained to appear on the user's profile, and cannot be placed on the

Functionality		Site		
		Facebook	Orkut	Twitter
Public messaging A space for the profile holder and contacts to leave notes: sometimes plaintext-only, sometimes richer.	Elements	Textbox, button, list of prior messages	Textbox, buttons, list of prior messages	Textbox, button, list of prior messages, recipient's username, @reply mechanism
	Effects	Quick/easy, communication (one-to-one), being overheard, social connectedness, anticipation of response, uncertainty	As Facebook, but not quite as easy (more buttons, can preview); added expressiveness, HTML-rich text, graphics	As Facebook, added load (need prior knowledge) but functionality is easily accessed
	Piecing Together	Audio messages recorded, situated on office door, play in loop	Similar to Facebook, but enable previews and video data	As Facebook, but shorter messages, situated semi-publicly in office reception (people must specify recipients' names)
Microblogs Succinct text updates, generally limited to 140 or 160 characters.	Elements	Textbox, button, list of prior posts, further options	Textbox, button	Textbox, button
	Effects	Fairly quick/easy, one-to-many communication, broadcasting info, presence online, openness, anticipation of response, uncertainty	Generally as Facebook	As Facebook, although even quicker/easier due to lack of additional options
	Piecing Together	T-shirt with scrolling text display of most recent post	As Facebook	As Facebook
Photo-sharing The ability to upload and caption photos, which can be commented upon	Elements	(These cells are merged because functionality is effectively identical across both sites) Complex photo upload process, option to annotate images, option to tag images, ability to browse photos		Not available
	Effects	Broadcasting information, sharing past experiences, presence online, openness, anticipation, reminiscence, uncertainty		
	Piecing Together	Novel tabletop displaying a sequence of photos		
Discussion groups Searching for and joining a group. Also called communities or forums: a mechanism for groups of like-minded people to make or maintain contact.	Elements	Textbox, button, list of results which links to group webpages	As Facebook, although list of results prioritises groups within the user's country	Textbox, button, use of #hashtag, list of matching tags and messages
	Effects	Quick, easy, insight into community availability, ability to join communities, potential connectedness and presence	As Facebook, although it is harder to gauge how active groups are, and to join them	Quick, easy, insight into community availability (and popularity) and activity, ability to join communities, potential connectedness/presence
	Piecing Together	Search a database about local community groups via web or smart-phone interfaces, including access to info on meeting regularity and popularity of groups	As Facebook	User specifies search concept, microprinter churns out matching Tweets

Table 4-1 Summary of application of TAPT to facets of social sites

profiles of others, but the richness of each mechanism's content is equivalent.

Twitter clearly stands apart from the other two platforms. It does not directly support photo-sharing, providing instead very lightweight mechanisms for public messaging, microblogging and discussion groups. All of these are achieved via the same textbox and button mechanism, provided on the homepage: public messaging and discussion groups occur via the use of Twitter-specific mechanisms, @replies and #hashtags. The analyses in Section 4.3.2 note that users must understand the @reply and #hashtag mechanisms, without specifying that users must understand how to use textboxes and buttons: this is because @replies and #hashtags are specific to Twitter, and not found elsewhere. By contrast, the other technologies which are mentioned in the analyses are commonly found on the web.

It is of note that the distribution of abstract and literal experienced effects varied across the analyses: for example, Teasing Apart public messaging and microblogging showed generally equal numbers of abstract and literal experienced effects, but Teasing Apart photo-sharing showed more abstract elements, while the analysis of discussion groups involved more literal elements. (These distributions held across platforms, but varied across functionalities.) This could be due to the nature of the functionalities considered: photo-sharing is an activity associated with feelings and impressions, and thus is perhaps a more 'abstract' experience than the activity of searching for a specific community online.

Having Teased Apart various aspects of functionality across different platforms, one can consider the overall social networking experience. Microblogging augments people's profiles, helping them equate to a physical presence by facilitating a sense of presence in the online community. Messages roughly equate to spoken conversations, held in public or private, and enhance communication and social connectedness. Photo-sharing is similar to seeing the moments in question as they happen, or sharing physical photos: like sharing printed photos, it involves reminiscence, openness and sharing of past experiences. Discussion groups appear to equate to active discussion forums, offering the potential for social connectedness and presence.

By including interactions between mutual friends and friends-of-friends, and by including text, images and discussion groups, these social websites begin to form a metaphor for village life. The sites instil a feeling of community and connectedness, not solely through direct chat, but through the combination of the above effects. A user on a social networking website won't generally just view a private message, or a photo, or a friend's status and comments on that status by friends and friends-of-friends, but will instead see a combination of microblogs, images, messages, upcoming events, etc. In a sense, they are seeing a personally-crafted, up-to-date bulletin board about their contacts, augmented with direct messages.

The functionality offered by social websites appears to map to various physical experiences such as conversation (direct or overheard), shared moments and seeing friends' activities. The overall effect is not unlike that of walking through a village populated by one's contacts, and observing and participating in activities and conversations.

Future work in this field could dig deeper into user awareness and understanding of social networking. For example, a public message thanking a friend for dinner is different from a private message with the same content: the former message broadcasts to friends of both parties that they had dinner together. How do experiences at each 'end' of a social interaction vary? That is, how do message senders and recipients experience the interaction? What about people who witness interactions between others?

4.4 Summary and Conclusion

This chapter presented an analysis of Dix's description of deconstruction and used it as inspiration for building a method, TAPT. This method was applied to several aspects of social networking websites (public messaging, microblogs, photo-sharing and discussion groups), in order to better understand these aspects and gain further insight into the TAPT process (leading to the addition of the steps described in Section 4.2.7 towards supporting the abstraction of experiences).

TAPT provided useful analytical information. Applying the method in a systematic, repeatable manner exposed emerging themes common across the items of functionality: for example, communication and sharing of material were common surface elements across the interactions that were considered. The public nature of each activity is of note: experienced effects across almost all items include being open or overheard, anticipation of responses and not knowing who beyond the intended recipient will see the content. Lack of certainty about one's audience also increases uncertainty regarding what manner of response will result, and when.

Experienced effects are a useful tool for distinguishing between superficially similar actions. For example, the analysis of public messaging and microblogging identified very different experienced effects of social connectedness (when messaging) compared to broadcasting information and consolidating one's online presence (when microblogging). As was hoped, TAPT was a useful tool for distinguishing between these two interactions which are similar on a superficial, surface level. Applying TAPT did not appear to highlight substantial differences between Facebook and Orkut, which suggests that either the sites are not deeply different or that TAPT is not sensitive to ways in which they differ.

It seems likely that the elements and effects associated with each item of functionality are more widely experienced when using social sites. For example, not all actions build one's presence online, but many do: consider microblogging, blogging, photo-sharing and completing a profile.

As noted in Section 4.3.3, the distribution of abstract and literal experienced effects varied across the analyses, according to the functionality being deconstructed. It seems likely that the nature of experiences directs the proportion of abstract and literal effects associated with them, but further work would be required to investigate this area.

It can be seen that Teasing Apart is useful for comparing multiple items in order to elicit commonalities and differences. It is of note that Teasing Apart is an inherently subjective process, requiring that the designer considers what 'surface elements' and 'experienced effects' make up an experience. One way to make the method more robust might be to include multiple practitioners in the process, meaning that outputs are not the result of a single, potentially biased perspective.

As well being useful for identifying themes, TAPT made re-imagining experiences far easier by abstracting those experiences. For example, it seems like a very big step to move from 'microblogging' to 'a scrolling display on a t-shirt'. By contrast, the steps from 'microblogging' to 'brief one-to-many communication' to 'a scrolling display' seem much smaller and more logical.

Further lessons learned are the importance of accounting for all aspects of an experience – that is, the experienced effects as well as the surface elements. It is straightforward to reason that a novel interface should offer a similar surface design to its web-based equivalent; however, it is less easy to account for the emotional implications of a transaction, such as the expectation of replies.

Mirroring the note in Section 4.1 concerning Dix's original experience deconstruction, TAPT does not represent a traditional engineering approach: it is an analytic and also creative tool, useful for understanding specific experiences and generating ideas for rebuilding these experiences. TAPT focuses on the semantic meanings of experiences rather than their syntactic presentation.

There is a parallel between TAPT and creative processes in general: as described in Section 3.1.1, creative processes involve a divergent stage and a convergent stage. Ideas are generated in the divergent stage (as with the abstraction of experiences in Teasing Apart and the brainstorming stage of Piecing Together) and evaluated and honed in the convergent stage (the rest of the Piecing Together).

The initial application of TAPT to facets of the social networking experience provided useful analytical insights into those experiences and showed that TAPT can yield useful results. However, successful use of TAPT by its creator only shows that it can be used by one person: further evaluation is required.

Existing literature suggests a perceived "pressure" in creativity and design research to seek quantitative results (Shneiderman, 2007). Similarly, Fallman (Fallman, 2010a) discusses a "common misunderstanding" involving assuming a close relationship between rigour and quantitative research. He argues that rigour is not necessarily related to the use of complex statistical methods, but instead is to do with the systematic nature of work and the clarity and transparency with which it is presented.

Shneiderman (Shneiderman, 2006) notes the use of multiple methods as a strong way to gain multiple perspectives and (Isomursu, 2007) discusses the same idea, remarking that the use of several methods can help 'triangulate' results and reduce problems should one method fail. The rest of this thesis presents a mixed methods evaluation of TAPT, as follows:

- practitioner evaluation of TAPT as compared against other methods (Chapter 5)
- expert review of design artefacts produced with TAPT (Chapter 6)
- case studies grounded in practice (Chapter 7)

Section 3.3 discusses the evaluation of creative processes and presents two frameworks for evaluating these (Kitchenham, 1996) (Furniss, 2008). Chapter 8 provides a holistic discussion of the methodology used in this research and weighs that methodology against the approaches described in Section 3.3.

The studies presented in this thesis primarily concern TAPT's use as a design tool, but some of the case studies in Chapter 7 involve applying TAPT in an analytical or evaluative sense.

This chapter has presented TAPT, a Software Engineering design process with a creative step, and demonstrated how it provides a systematic approach for understanding and replicating experiences.

Chapter 5 Comparative Evaluation

This chapter presents the methodology and results from a comparative evaluation of TAPT against two other approaches to designing novel computer systems (Scenarios, chosen as representative of current practice in industry, and Unstructured Discussion, chosen as a neutral baseline). The comparative evaluation was followed by an expert review (described in Chapter 6) and a number of longer studies which are presented in Chapter 7 as case studies.

The comparative evaluation was conducted to gain more objective insights into TAPT's efficacy as a tool for understanding and replicating experiences. The initial study described in Section 4.3 was conducted by the author and therefore subject to bias, while results from other practitioners would be more objective. Additionally, recruiting a larger number of participants for the second evaluation led to more results than could be gained from just one researcher, such that some results were statistically significant.

The study was preceded by a series of small pilot studies, conducted to refine the description of the TAPT process. 43 people took part in the comparative evaluation, forming 21 groups. 33 participants were from IBM, and 10 from the University of Southampton. The study ran over 5 separate sessions between 30th November and 10th December 2009, and each session had between 6 and 12 participants.

Carrying out the formalised TAPT process resulted in clear, tangible output. The TAPT forms act as design artefacts, recording observations and decisions made and providing a simple audit trail of the creative process. Figure 5-1 shows one such form, in which a group applied TAPT to a design task involving wikis.

Teasing apart results

Teasing apart ____________				
Experience (1)	**Surface elements (2)**	**Experienced effects (3 & 4)**		**Distilled experience (5)**
		Literal (3.1)	**Abstract (3.2)**	
Wikipedia allows users to search and find pages about their topic, read info + pictures, follow links for more data add and modify pages and review how and why change made	View Text + pictures follow links between pages, see change log, uncontrolled modifications unknown audience	See text See pictures Slowly changes with other people edit, Harder to see old information Easy to make changes Finding related info	sharing information learning, expecting to find some info. Uncertainty of amount to find Unease with accuracy or possible offensive comments Satisfaction sharing adding / correcting information	Learning about a topic I choose with suggestions to find more, uncertainty of accuracy and depth Ability to give others the benefit of my knowledge in semi permanent way
		Remember to note key effects!		

Piecing together results

Piecing together ____________	
Distilled experience	Learning about a topic I choose with suggestions to find more Uncertainty of accuracy and depth Ability to give others the benefit of my knowledge in semi permanent way
Context	Physical area of museum, touch map, video screens video camera + sound recorder
Description	Touch map to select user stories with location. Touch time line to select when View / listen video reminiscences Rerecord over other people / add new Voice recognition to get links – "see Brandenburg gate"

Figure 5-1 Teasing Apart a wiki and Piecing it Together in a museum context

Each group produced two artefacts, where an ‘artefact’ is a design output from one of the three given methods. TAPT was used 17 times, Scenarios 12 times, and Unstructured Discussion 13 times.

Outputs of the study were:

- Questionnaire responses
- Audio recordings of five group discussions (one from each session)
- 42 design artefacts:
  - 14 microblog artefacts
  - 15 picnic artefacts
  - 13 wiki artefacts

The resultant data is primarily qualitative in nature, although the Likert-scaled questions yielded some amount of quantitative data, as did a numerical analysis of the artefacts.

Section 5.1 describes the pilot studies, before 5.2 details the experimental design. Section 5.3 provides an analysis of design artefacts, and Section 5.4 discusses participants’ responses to the questionnaire. Finally, Section 5.5 offers some conclusions.

5.1 Preliminary Work

Before testing TAPT on a larger scale, it was necessary to conduct a number of small pilot studies in order to facilitate iterative improvements to the TAPT method. These seven pilots were conducted in a commercial work environment and a university environment.

Participants were selected individually and opportunistically: the criteria were that they were professional Software Engineers able to spare up to an hour of their time to walk through the TAPT process and provide constructive feedback. Participants were Software Engineers (at IBM) and late-stage PhD students (at the university).

Participants were asked to apply TAPT to transform any experience of their choice, as long as either the starting or ending context was technological. If a participant had no ideas, the researcher suggested experiences such as writing email or reading a stream of microblog updates.

The process for six of the seven pilot studies was that the individual participant would sit down with the researcher and apply TAPT to an experience, giving verbal feedback during the process using a simple 'think-aloud' protocol (Johnson, 1994). At the end, the researcher and participant would have an informal discussion about the participant's experience using TAPT.

The exception to this process was the final pilot, which was run with a group of three final-year PhD students, who were asked to apply the Scenarios method to a task (redesigning wikis for a museum). In this pilot, participants were also asked to fill in an opening questionnaire, a task questionnaire after the exercise, and a closing questionnaire. The intent of this final pilot was to gather data on how much time would be required by small groups working on the tasks and to gain feedback on the questionnaire design. The final pilot resulted in feedback on the presentation of the Scenarios method, resulting in a clarification about personas and expected output, and a simplification of the method description.

This series of pilots resulted in many changes to the presentation of TAPT. At the beginning, TAPT was described in a step-by-step list where the corresponding example was given alongside each step, in italics. For example:

1. **Experience:** Brief (maximum 200-word) description of the chosen functionality and the experience of using it. *(For example: The 'wall' on Facebook is a space associated with a user's profile page, where friends and the profile-holder can share a limited quantity of text.)*
2. **Surface elements:** These are generally nouns ('line', 'box', 'arrangement of photos') and adjectives ('bold', 'simple', 'complex') relating to the design. *(A box for up to 1000 characters of plaintext; a 'share' button; a list of previous messages, most recent first.)*

By the end, the TAPT instructions and examples were separated and presented as a table (for Teasing Apart) and a numbered list (for Piecing Together), shown in Table 5-1 and below the table. A table layout was chosen for Teasing Apart as feedback suggested people liked having a blank table to fill in, and that presenting the instructions in a table of the same format made it easier to understand which material was to be placed where. The numbers in the table show the order in which to carry out each step.

<table>
<tr><th colspan="5">Description of Teasing Apart, step by step</th></tr>
<tr><td rowspan="3">Experience (1) Brief description of the chosen functionality and the experience of using it.</td><td rowspan="3">Surface elements (2) These are generally nouns (‘line’, ‘box’, ‘arrangement of photos’) and adjectives (‘bold’, ‘simple’, ‘complex’) relating to the design.</td><td colspan="2">Experienced effects
These focus on the physical, emotional and intellectual effect upon participants, and tend to be abstract nouns (‘excitement’), noun/verb pairs (‘hunger sated’) and perhaps adverbs (‘quickly’). There are two types of effect, shown below…</td><td rowspan="3">Distilled experience (5):
Consider your table of information, particularly the aspects which you think are key to the experience, and use it to describe the experience as a sentence. Try to keep your sentence neutral: for example, you might mention ‘broadcasting’ information rather than ‘showing’ it, because ‘showing’ implies a visual broadcast.</td></tr>
<tr><td>Literal (3.1) Concrete results such as ‘a loud noise’, ‘broadcast information’.</td><td>Abstract (3.2)
Relating to emotional and intellectual effects, such as ‘excitement’, ‘co-experience’. This step is important: dig deep!</td></tr>
<tr><td colspan="2">(Step 4) Review the above lists of effects, and identify effects that seem especially important, unique or key to the experience. Underline them.</td></tr>
</table>

Table 5-1 TAPT instructions in a table format

The steps of Piecing Together:

1) *Brainstorming, particularly using the key effects identified earlier. Feel free to use scrap paper, whiteboards etc. if this is helpful. If you find it hard to come up with ideas, consider things you might change from the original experience. These might include:*
   - *the modality of communication (vocal, textual, musical, graphical, symbolic)*
   - *the technology used (pen and paper, PDA, telephone, cardboard)*
   - *scale (are we working with one person? Tens? Hundreds? How big is the physical space? How large are the items with which people interact?)*

   You might also want to think about what technologies traditional occur in the original and the new environments – office might traditionally include telephones, whiteboards and desktop computers, while parks tend to traditionally have benches, maps printed on large boards and fountains. What can you use?
2) *Scenario building. Using your ideas, build an example reconstruction. Below are two examples:*
3) *Check your reconstruction:*
   i. *Considering the distilled experience and the key elements of the original experience, have you included everything you want to? (It may be that you choose to drop some key effects from the original experience. This is fine.)*
   ii. *Have you incorporated any new key effects which you did not intend?*
   iii. *Refine the pieced together scenario accordingly.*
   iv. *Repeat the above three steps until you are happy with the scenario.*

The main changes made to the instructions were:

- Separating the Teasing Apart and the Piecing Together instructions.
- Providing instructions in a tabular format.
- Separating examples of TAPT's use from its description.
- Providing two examples of TAPT's use (Christmas crackers and microblogging), demonstrating its use for physical-to-digital and digital-to-physical transitions.
- Providing a blank table for participants to fill in (this also provided a cue about how much text might be expected, which also meant the text prompting the initial experience description did not need to specify 200 words as a maximum length).
- Including a prompt to underline key effects.
- Numbering cells in the TAPT table to link them with each step in the instructions.

5.2 Experimental Design

After the pilots were completed, the comparative evaluation was conducted. This section describes the design of the evaluation. After an overview, this section describes recruitment, ethical aspects and experiment materials (Section 5.2.2), tasks and methods (Section 5.2.3), the allocation of these (Section 5.2.4) and questionnaire design (Section 5.2.5). It then goes on to describe the approach taken to numeric analysis of questionnaire responses (Section 5.2.6) and TAPT artefacts (Section 5.2.7), the impact of participant background and task allocation (Section 5.2.8), how t-tests were used once data had been gathered (Section 5.2.9) and finally unexpected uses of TAPT (Section 5.2.10).

5.2.1 Overview

Participants were placed in groups of two to three (based on the typical size of groups conducting design work in IBM[19]). Having individually filled in a preliminary questionnaire, participants were asked to work in their groups to redesign experiences in new contexts, using one of three approaches:

1. Unstructured Discussion
2. Scenarios
3. TAPT

[19] Based on discussion with user experience professionals at IBM.

Unstructured Discussion (UD) was included as a neutral baseline, while Scenarios was selected as representative of current practice in industry.

Each group was asked to carry out two design tasks during the time, with each participant completing a written questionnaire after each task. Each task was carried out using a different technique, selected from the three listed above. Once all groups had completed their two exercises, individuals would fill in a final, comparative questionnaire. A group discussion and debrief was held for the remaining time, which was subject to audio recording.

The procedure took 2.5 hours:

0:00 - 0:15	Investigator explains procedure and answers any initial questions. Participants fill in preliminary questionnaire
0:15 - 1:00	Groups begin first re-imagining, and individually complete questionnaire on this
1:00 - 1:45	Groups begin second re-imagining, and individually complete questionnaire on this
1:45 - 2:00	Participants complete final, comparative questionnaire
2:00 - 2:30	Discussion and debrief.

The discussion session was an opportunity for participants to explore the value of the methods used, and to ask any questions which had arisen during the session. Discussions were recorded and anonymised. The goal of the discussions was to allow participants to highlight key points which they felt were important. The investigator launched these sessions by asking whether any participants felt that they had written something important when responding to questionnaires, which they wished to share with the group. An effort was made to keep the discussion moving, and to focus on methods used rather than resultant designs.

5.2.2 Recruitment, ethical considerations, and experiment materials

The investigator gave an introductory half-hour talk about her research at both IBM and the university, which resulted in recruitment of volunteers.

Forty-three Software Engineers participated in the study. The majority (33) worked at the IBM Hursley Laboratory in Hampshire, while seven were employees and three were PhD students at the University of Southampton. Most participants were British citizens.

It was not possible to run a blind trial (where the participants were unaware of the novel method) because of the necessity of using experienced Software Engineers who would already be aware of existing methods, and be easily able to spot any new approach.

The only identified risk with the experiment was the extent to which it intruded on participants' time, as it lasted 2.5 hours. This length of time was necessary to obtain adequate data: running several shorter sessions would lead to variation in participants, moods and approaches. To mitigate the risk, participants were made aware that they could leave at any time and were provided with light refreshments.

Various written materials were given to participants at the opening of the experiment. These included a participant information sheet, a consent form, and a more detailed outline of the plan for the session, including timings. Examples of these documents are given in Appendix E: Comparative Evaluation Materials.

5.2.3 Tasks and methods

Each group carried out two design tasks during the study, responding to two of three possible tasks, and addressing those two tasks with two of three possible methods. Methods tested were TAPT, Scenarios, and Unstructured Discussion.

The Scenarios method was selected on the basis of discussions with usability practitioners at IBM: in addition to Scenarios, these discussions touched upon the use of storyboarding, heuristic evaluation, and conceptual modelling. Practitioners confirmed that Scenarios was the most appropriate choice because it is representative of current usability practice.

Finally, Unstructured Discussion (UD) was included: UD was created for the purpose of the experiment. It would never be used in practice, but was incorporated as a neutral baseline, against which the effects of the other two methods could be gauged.

The three methods, as described to participants, are fully presented in Appendix E: Comparative Evaluation Materials. A brief summary is given here:

Participants using **TAPT** were given detailed instructions about the process, and two examples of possible TAPT output: designing Christmas crackers on a webpage, and photo-sharing in the home. Two examples were given, because feedback from the pilot studies suggested that being able to see TAPT's use in both physical and digital contexts was helpful.

Scenarios (also known as Scenarios and Personas) was described as "a tool for imagining ideal user interactions". Participants using this method were given a description which boiled down, briefly, to these steps, drawn from Cooper's descriptions of Scenarios (Cooper, 2007):

1. Brainstorm around what it is you need to design
2. Consider types of user of the system you're designing (personas), and create the persona(s) which seem key to the system
3. Construct high-level scenarios of system use, from the users' perspectives.

The steps were followed up with an example persona and scenario, also taken from Cooper's material.

Unstructured Discussion was presented as a 'group discussion', and described thus:

> *A very informal method, this simply involves carrying out a verbal discussion within your group about how to solve the problem at hand. Please do not write or draw ideas during your discussion, but write a paragraph describing your envisioned design once you have decided upon its details.*

Regardless of method used, all participants were asked to write one paragraph describing their envisioned design, to be handed in to the investigator upon completion (for TAPT users, this was the 'description' part of the TAPT table).

As well as design descriptions, participants were asked to provide the researcher with all written materials generated during the study, including rough notes, TAPT tables, Scenarios and Personas.

Participants using TAPT or Scenarios were allowed to make freeform notes if they so desired, but those using Unstructured Discussion were not. Making freeform notes is something which happens in practice in industry, but Unstructured Discussion was not supposed to be realistic: it was a baseline for measuring the impact of the other approaches.

The task descriptions were:

1. Original experience: picnicking with friends
   New experience: A company which sells picnic products have asked you to design a website which re-provides the experience of enjoying a picnic with friends. They have plentiful funding and it is clear that (if it is helpful for achieving your goal) then [sic] you are welcome to apply technology in novel ways
2. Original experience: Reading and extending a wiki
   New experience: A museum about the Berlin Wall have [sic] been awarded funding for an extension. They want to create an area where members of the public can provide comment: the aim is that contributors will share their experiences, thoughts and feelings about the Berlin Wall. The museum is keen to use technology to encourage people of all ages and backgrounds to contribute and want to replicate the feel of Wikipedia
3. Original experience: reading, writing and commenting on status updates on social networking sites
   New experience: A care home for the elderly have observed interest from their residents in strengthening their community. Inspired by social networking sites such as Facebook, they are interested in building a system to enable residents to provide short ‘status updates’, alongside the ability to access and comment upon updates from other residents. The system will be used only by old people, living in the home: it caters for a relatively small network of people. The care home ask you to design this system.

The above tasks were respectively chosen as examples of re-providing a physical social experience on the web; re-providing a web-based experience in a public physical context; and re-providing a web-based experience using pervasive technologies in a context where accessibility is of paramount importance.

The researcher chose to avoid straightforward examples such as ‘sending a greetings card’ or ‘receiving a letter’. This is because the focus of inquiry was how to redesign experiences for which a straightforward transition is less obvious. Additionally, these alternative examples have already been addressed in the digital context (for example, consider e-cards and email), and pre-existing digital solutions to the tasks would be leading for participants.

The investigator checked at the beginning of each exercise that groups were familiar with the background for their tasks: that is, that those working on microblogging and wikis were aware of the relevant phenomenon, and also that participants from a non-British background were familiar with the concept of picnicking. In the event, all participants were comfortable with their tasks, but had this not been the case, the investigator would have rearranged groupings to ensure that this was the case.

5.2.4 Task and method allocation, impact of participant background

Twenty-four groups of two to three participants in size participated in the experiment. The same method or task was not allocated to groups more than once across the two exercises, because prior experience with either of these during the second exercise would affect the results. To combat any influence from, for example, the first method used upon the second method used, groups were allocated to equally distribute the method order. For example, an equal number of groups applied TAPT followed by Unstructured Discussion (where prior use of TAPT may influence application of Unstructured Discussion) as applied Unstructured Discussion followed by TAPT (where the Unstructured Discussion approach may influence interpretation of TAPT). Similarly, the order of tasks varied across groups.

To illustrate this, Table 5-2 shows the group, task and method allocation of the first five participants:

Participant	**Group**	**First method and task**	**Second method and task**
1	A	TAPT Wiki	Scenarios Picnic
2			
3	B	TAPT Microblog	UD Wiki
4			
5	C	Scenarios Microblog	TAPT Picnic

Table 5-2 Allocation of groups, tasks and methods

Table 5-3 shows the number of groups to respond to each possible sequence of tasks, and Table 5-4 shows the number of groups to respond to each possible sequence of methods.

First task	Second task	Number of groups to complete tasks in this order
Microblog	Picnic	4
Microblog	Wiki	3
Picnic	Microblog	4
Picnic	Wiki	3
Wiki	Microblog	3
Wiki	Picnic	4

Table 5-3 Number of groups allocated each sequence of tasks

First method	Second method	Number of groups to use methods in this order
UD	Scenarios	3
UD	TAPT	4
Scenarios	UD	2
Scenarios	TAPT	4
TAPT	UD	4
TAPT	Scenarios	3
TAPT	TAPT	1

Table 5-4 Number of groups allocated each sequence of methods

Note that one group (Group I) applied TAPT to both of their tasks, erroneously. Their second design artefact has been included in analysis of artefacts (where we are interested in the impact of using TAPT upon the content and ideas generated), but their questionnaire responses from the second use of TAPT were discarded (as their answers to questions about, for example, the ease of use of the method will have been influenced by their prior experiences).

5.2.5 *Questionnaire design*

Questionnaires were chosen as the primary method for eliciting participant responses to the methods used. Although interviews would enable the researcher to acquire more in-depth answers, the scale and set-up of this particular study precluded this: it was important to obtain information on their opinions of the methods immediately after the time of use. With six to twelve participants carrying out the study at any time, interviews would not have been able to achieve this goal. The additional use of a group discussion at the end of each session enabled participants to voice strongly-held opinions, and was complementary to the use of questionnaires.

Three questionnaires were used in the study:

1. An opening questionnaire, intended to elicit participants' previous experience of software design processes (Table 5-5).
2. A questionnaire about the design task, administered twice during the study, after completing each design exercise (Table 5-10).
3. A closing questionnaire, to elicit which of the two methods was felt to be the easiest to use, most useful, and most effective at replicating experiences (Table 5-11).

The primary aim of the questionnaires was to find out whether the methods were useful, and whether the methods helped participants replicate the given experiences.

Question	**Purpose**
What is your previous experience of software design processes? (Consider, for example, scenario building, storyboarding and UML.) Do you perceive these processes as generally positive or negative?	Elicit participants' level of experience and perceptions of software design approaches (to contextualise results).
What participant number are you?	Track participant.

Table 5-5 Opening questionnaire (responses as freeform text)

<table>
<tr><th>Question</th><th colspan="2">Purpose</th></tr>
<tr><td>What is your participant number, which task did you work on, which method did you use?</td><td colspan="2">Track participant, task, and whether it was the first or second task.</td></tr>
<tr><td>Were you able to carry out the task? If not, why?</td><td rowspan="3">Elicit whether the task was explained well.</td><td>Whether they were able to carry it out.</td></tr>
<tr><td>Were the task and the method explained clearly? If not, how so?</td><td>Whether it was presented clearly.</td></tr>
<tr><td>Did you find any steps of the process difficult? If so, which ones, and how so?</td><td>Whether any steps were difficult.</td></tr>
<tr><td>Do you think there are other methods that you might have used to do this sort of work? (Ignoring any which you have used during this study.) Please explain why they are better or worse than the one you applied.</td><td colspan="2">Elicit whether alternative methods are perceived as better.</td></tr>
</table>

Table 5-6 Design task questionnaire: initial questions (responses as freeform text)

Question	Purpose
How well did the method do at improving your understanding of the original experience? How so?	Elicit usefulness for increasing understanding of underlying meanings.
How well did the method do at improving your awareness of emotional effects of the original experience? Why?	Elicit usefulness for increasing understanding of underlying meanings (emotional effects).
How well did the method improve your understanding of the social context of the original experience? Why?	Elicit usefulness for increasing understanding of underlying meanings (social context / effects).
How well did the method improve your understanding of the changing perceptions, expectations and reactions of users before, during and after the experience you analysed? In what way?	Elicit usefulness for increasing understanding of underlying meanings (the sequence of the interaction).

Table 5-7 Design task questionnaire: questions on usefulness

(responses on a 5-point Likert scale with space for free text)

Question	Purpose
How well does your artefact recreate the experience of the original? Does it support experiences analogous to the original?	Elicit usefulness for helping replicate underlying meanings.
How well does your redesigned experience replicate the emotional aspects of the original? Would users experience similar emotional reaction as with the original?	Elicit usefulness for helping replicate underlying meanings (emotional effects).
How well does your design replicate the social context of the initial experience? In what way?	Elicit usefulness for helping replicate underlying meanings (social context / effects).
How well will user perceptions, expectations and reactions around the new experience match those of the original? Why?	Elicit usefulness for helping replicate underlying meanings (the sequence of the interaction).

Table 5-8 Design task questionnaire: questions on replication

(responses on a 5-point Likert scale with space for freeform text)

Question	Purpose
How well did the method scaffold a discussion about the problem?	Elicit usefulness in providing a vocabulary for discussion.
How well did the method support you in creating imaginative and novel designs?	Elicit usefulness in fostering creativity and enabling lateral thinking.
How well are the artefacts you generated (written text, tables, concept maps, scenarios) suited to use in the workplace? (I.e., for documentation or communication.)	Elicit usefulness in leading to usable artefacts.
How well did you replicate the surface design elements of the original experience? (These are tangible design elements, such as textboxes and graphics in a webpage or shiny wrapping paper and a paper hat in a Christmas cracker.)	Elicit usefulness in replicating superficial aspects of the starting experience.

Table 5-9 Design task questionnaire: miscellaneous scaled questions

(responses on a 5-point Likert scale with space for freeform text)

Question	Purpose
Some of the approaches in this study were formal (Scenarios and TAPT), and some less so. What was the effect of this structure or lack of structure on the way you worked?	Elicit usefulness in providing a structured / systematic approach.
Considering the method as a whole, do you have any other comments about its usefulness or value?	A catch-all to capture any other reactions to the study.

Table 5-10 Design task questionnaire: final questions (responses as freeform text)

The three comparison questions were given in the closing questionnaire, designed to elicit the relative ease of use, usefulness, and effectiveness at replication. They are shown in Table 5-11.

Question	Purpose
Considering the two methods you have applied today, which did you find easier to use? (Consider how straightforward the methods were, and how long they took to apply.) Why?	Elicit methods' relative ease of use.
Which did you feel produced more useful (powerful, practical, relevant) results? Why?	Elicit methods' relative usefulness.
Which did you feel was more effective at replicating the experience you were considering? Why?	Elicit methods' relative effectiveness at replication.
What participant number are you?	Track participant.

Table 5-11 Closing questionnaire (responses as freeform text)

Full copies of the questionnaires can be seen in Appendix E: Comparative Evaluation Materials.

Table 5-7 and Table 5-8 comprise the bulk of the main questionnaire, which was presented to participants as a table of eight central questions about the method's usefulness for a) increasing understanding and b) supporting replication of four aspects: the overall experience, emotional aspects, the social context, and the event as a temporal sequence (i.e. consideration of expectations beforehand and responses afterwards). The questionnaire provided further prompts for the questions about emotional, social and temporal aspects:

- Emotional aspects. For example, these might include feelings of diligence, happiness or anticipation
- The social context and social effects of the experience. For example, people's social expectations beforehand and social fulfilment afterwards
- User perceptions, expectations and reactions during the sequence of events.

These aspects were given particular attention for reasons that arose in Section 4.2.1: prior literature as well as the TAPT method refers to emotional and social aspects of design. The third question, concerning perceptions, expectations and reactions before, during and after the experience is rooted in the work of Wright et al (Wright, 2003), who describe experience as four threads, one of which is spatio-temporal, that is, the experience as situated in space and time.

Participants of the evaluation were asked how well their methods did at helping them understand and replicate emotional, social and sequential aspects based on these contributions.

The study itself ran over the course of five separate sessions, four at IBM Hursley and one at the University of Southampton. After the first session, three minor changes were made to the questionnaire, for clarity. These were:

1. Changing the word 'process' to 'method' in several questions
2. Adding the sentence '(Ignoring any which you have used during this study.)' to the following question: 'Do you think there are other methods that you might have used to do this sort of work? Please explain why they are better or worse than the one you applied.'
3. When asking which task was completed and which method was used, instead of having the numbers 1, 2 and 3 next to tick boxes, using prose (picnic, wiki, social networking; Scenarios, Group Discussion, TAPT).

5.2.6 Approach to numeric analysis of questionnaire responses

Some questionnaire responses were formatted as a five-point Likert scale. Participants were asked to rate how well the method they used did, from 'very badly' through 'badly', 'ok' and 'well' to 'very well'. Responses could therefore be scaled to a numeric rating, from 1 (very badly) to 5 (very well). Items such as mean and variance could be calculated from these numbers, which were used towards conducting t-tests. T-tests were chosen because they offer a simple, transparent way to test the statistical significance of results.

When calculating t-tests, it is necessary to choose an alpha level. The alpha level represents the possibility of the observed data occuring at random if the hypothesis is wrong: so, for example, an alpha level of p=0.05 gives a 95% probability that the hypothesis is correct. This thesis uses an alpha level of p=0.05.

When using an alpha level of p=0.05, one may expect that one test in twenty will be statistically signifiant by chance. In this work, 36 pre-planned tests were conducted (some additional tests were conducted post-hoc: these are discussed later). One could expect 2 of these tests to yield significant results purely by chance. As will be seen, in the event 11 of the tests gave significant results.

A disadvantage of carrying out numerous t-tests on one dataset is that there is a danger of finding significant effects by chance ('fishing' for results). Arguably, the alpha value (the boundary of statistical significance) for such results should be considerably lower than the standard 0.05.

An alternative to t-tests would be a more complex statistical procedure such as the Analysis of Variance (ANOVA), which indicates whether there is a 'main effect' and whether there are 'interactions' (for example, between a group such as 'users of TAPT' and a response category such as 'usefulness'). The ANOVA only uses 'post-hoc' t-tests, to determine where any effects come from: as these t-tests are planned, they cannot be criticised on grounds of 'cherry-picking' (Troscianko, 2010).

The advantage of the ANOVA is that it uses the variance in the whole dataset to estimate the reliability of the results, while the multiple t-tests will concern only the two columns of data being tested at any one time.

It can, however, be argued that interpreting results becomes more difficult when more complex statistical operations are performed. T-tests were chosen in order to keep results as simple and transparent as possible.

5.2.7 Approach to numeric analysis of TAPT artefacts

For each of the three tasks, the resultant TAPT-generated artefacts were collated: there were five microblog, six picnic and six wiki artefacts. Data from each step of the analyses were grouped, so that initial experience descriptions were side-by-side, as were surface elements, literal effects, and the other analysis components.

The analysis involved identifying the number of elements and effects to be identified, taking a greater number of elements and effects to correlate with greater creativity.

The author began by noting the length and content of initial experience descriptions.

Next, surface elements were considered. A normalised table of elements was produced by coding the data. The aim was to identify synonyms across groups (for example, 'status' and 'status box', or 'textbox' and 'box to type in'). The author then counted the number of elements produced by each group[20]. An alphabetised list of every surface element per task was made, where repeated items were marked in bold, and annotated with a number in brackets showing their frequency of use. For example:

1. Authentication
2. Brief
3. **Broadcast (2)**

Next, literal effects were dealt with. As with surface elements, the number of effects was noted. Additionally, the number of key effects (effects chosen as essential to the experience) was noted. The lists of effects were normalised, and two alphabetised lists were generated (again, with repeated items marked as described for surface elements): the second list included key elements only. The same process was conducted on abstract effects, again giving separate attention to key effects.

[20] Note: this affected the count. For example, Group Q listed picnic items including 'snack eggs', 'mini-pastry', 'cocktail sausages', 'quiches' and '6 little pork pies': no other group took such a detailed approach of listing individual food items. The author normalised this set of items into one item, 'food', which was used when counting elements. (Conversely, Group O listed one item called 'food/drink': this was separated into two items during normalisation, to reflect the approach taken by the other picnic groups.) This is because differences in the level of granularity to which groups work are not of primary concern here: instead, we are interested in how many meaningful aspects of an experience TAPT helped participants to identify.

When in doubt, the author did not separate or group items: for example, Group P listed three literal effects of 'Personal reply', 'Provoke discussion' and 'ask for help': all of these could be argued as one element, 'conversation', but they seemed sufficiently different to be left as separate units.

The distilled experience description was treated as per the initial experience description, with notes made of length and any striking content.

The Piecing Together phase was also considered. The chosen context of the new experience was noted (for example: of the five microblog analyses, two described the new context in terms of care homes, and three talked about specific devices). Finally, the experience descriptions were considered, once again in terms of length and any interesting facets. Extraneous material was excluded: for example, one wiki response listed new effects and an extension (all written outside the TAPT table, which had been completed). This data was excluded from the numbers used here, on the basis that the numbers were intended to measure the straightforward use of methods and additional materials distort that measurement. (This additional material was subject to later discussion: see Section 6.3.2.)

A methodological note: some tables include percentages, such as the percentage of a group's surface elements to have been chosen by another group answering the same task (see Table F-5 in Appendix F: Numeric Data on Comparative Evaluation TAPT Artefacts). The percentages are calculated by taking the average of each individual group's percentage of elements repeated by other groups. In the microblog example, this is the average of 30, 42.85, 50, 62.5 and 33.33: this gives 43.74%. An alternative approach would have been to divide the average number of repeated elements (or effects) by the average number of elements (or effects) and multiplying by 100, which would give an average percentage of microblog elements repeated by other groups of 43.59%, not 43.74%, as given in Table F-5.

This first method was chosen over using the aggregate averages for groups because it is less perturbed by outliers. For example, if one group responding to a task included many more elements than the others and chose a wildly different percentage of elements (or effects), it would skew a percentage calculated by the simpler approach. In contrast, the described approach can cope with this scenario.

Appendix F: Numeric Data on Comparative Evaluation TAPT Artefacts gives the full set of numbers used.

5.2.8 Impact of participant background and task allocation

Inevitably, prior expertise influenced responses. For example, Participant 26 answered 'ok' to the question on how well TAPT helped improve understanding of wikis, commenting “Having worked with many teams on developing wikis, I was already familiar with the experience”.

Table 5-12 shows average responses to two key questions on usefulness and replication (on a scale of 1 to 5, where 1 is ‘very bad’ and 5 is ‘very good’), sorted according to whether the answers were provided after the first or the second task. The questions considered are:

1. How well did the method do at improving your understanding of the original experience? How so?
2. How well does your artefact recreate the experience of the original? Does it support experiences analogous to the original?

	TAPT		Scenarios		UD	
	1st task	2nd task	1st task	2nd task	1st task	2nd task
Understanding	3.7	4.3	3.2	3.9	3.9	2.9
Replication	3.6	3.6	3.7	3.5	3.8	3.5

Table 5-12 Responses to questions after participants’ first and second task

As can be seen, both TAPT and Scenarios users rated their methods more highly at supporting understanding after having used UD (by 0.6 and 0.7 respectively), which by contrast was rated less well when used after one of the other two methods (score reduced by 1.0). This decreased opinion of UD as seen after a structured method and increased opinion of the other methods when seen after UD suggests that UD did not fare well at supporting understanding when compared to the other two methods.

Ratings of methods for replication stayed approximately the same over first and second use of methods, bar a small drop in the rating of Scenarios and UD (0.2 and 0.3 respectively) when it was their second usage. This could suggest that those two methods compared less well to TAPT, but the difference in ratings is rather small.

Another question is whether IBM and university participants differed significantly in their responses. Table 5-13 shows average responses to the same two questions. As can be seen, these are largely similar, although IBM participants rated TAPT as much better at replication than university participants. Relatively few of the ten university participants used TAPT (the majority were allocated Scenarios and UD), meaning it is impossible to tell if this result is significant.

	TAPT		Scenarios		UD	
	IBM	Uni	IBM	Uni	IBM	Uni
Understanding	4.0	4.5	3.3	3.9	3.4	3.4
Replication	3.7	2.5	3.6	3.6	3.6	3.7

Table 5-13 Responses to questions (by participant background)

5.2.9 T-tests

Section 5.2.6 discusses the strengths and weaknesses of t-tests. The following statistically significant results were found with t-tests which were planned from the outset, and are therefore the most robust:

- TAPT was more useful overall than Scenarios and Unstructured Discussion (Table 5-15)
- TAPT was more useful for understanding emotional aspects than Scenarios and Unstructured Discussion (Table 5-16)
- TAPT and Scenarios were more useful than Unstructured Discussion for understanding perceptions and expectations (Table 5-20)
- TAPT was better at replicating expectations and perceptions than UD (Table 5-25)
- TAPT and Scenarios scaffolded conversation better than UD (Table 5-27)
- TAPT and Scenarios were more suitable for workplace use than Unstructured Discussion (Table 5-30)

Table 5-18 shows TAPT and Scenarios' ratings to be statistically better than those of UD for increasing awareness of social effects, but this test involved examining a subset of the pool of participants (participants who were conducting their second task). Although the results shown in Table 5-18 are statistically significant, they should be regarded as less robust in this broader context. The data in Table 5-18 has an alpha level of 0.005 for the TAPT > UD comparison (99.5% probability of correctness) and of 0.025 for the TAPT > Scenarios comparison (97.5% probability of correctness).

Concerning replication of changing expectations and perceptions, TAPT and Scenarios again did similarly well, followed by Unstructured Discussion: TAPT and Scenarios were never rated as doing 'Very Badly' at this. TAPT was statistically significantly better at this facet than Unstructured Discussion.

5.2.10 Unexpected uses of TAPT

Participants did not always use TAPT as expected. For example, the researcher expected starting experience descriptions to concern microblogging, picnicking and wikis for the three tasks. Two of the five microblog groups, however, discussed care homes (rather than microblogging), while one wiki group created a list of functionality rather than prose to describe an experience. Similarly, three of the five microblogging groups gave a technology-oriented context description (“Handheld devices (individual)”; “Interactive TV”; “Mobile Device (TABLET)”) rather than a description oriented around a care home.

Some TAPT analyses had surprising content: for example, Group P listed very specific abstract microblog effects, including ‘disappointment’, ‘excitement’ and ‘surprise’: it seems unlikely that they thought all of these happen every time someone microblogs. Perhaps this group was considering a breadth of possible experiences from multiple instances of microblogging, rather than one instance of microblogging. This ties in with material from the expert review, where a reviewer felt that picnics are too diverse as social events to be abstracted into one description. (See Section 6.5.1.)

Participants did not always seem to be aware that dropping certain functionality from the starting domain was not necessarily a bad thing: particularly, most of the designs produced in the wiki task enabled museum visitors to edit one another’s shared opinions. See Section 5.3.3 for more on this. An improvement to the presentation of the TAPT process would therefore be to emphasise that it is acceptable for practitioners to choose to drop aspects of the starting experience, if this seems appropriate.

5.3 Analysis of Resultant Artefacts

This section presents and discusses the design artefacts resulting from the experiment, considering in turn microblog, picnic and wiki artefacts. Appendix G: Artefacts from the Comparative Evaluation) presents two sample artefacts and summarises all resultant artefacts from this experiment.

5.3.1 Microblogging

Five groups used TAPT in response to the microblogging task: two used handheld devices for voice or visual interactions (one of these included staff assistance), one used a menu-based TV interface, one used a web-interface, and one used a tablet interface. An example artefact is:

A wireless touch screen tablet with a very simple UI. Display consists of a few large buttons for simple functions. Eg. "Post Status," "Read," "Comment" etc. These buttons are contextual. Device allow [sic] text-to-speech output. Each user has their own device so authentication is not required. Status can include pre-canned sentences. These are global to the community and can be added by any capable user. Status can be added automatically (RFID Positioning etc). Text input via on-screen keyboard.

Microblog groups generally identified literal effects such as information and sharing, while abstract effects included connections, sharing and anticipation. Distilled descriptions focused upon community, sharing and staying in touch. Implementations were diverse, varying from voice to tablet interfaces.

Microblog participants seemed to agree less about what the microblogging experience is than participants addressing other tasks. For example, on average only 44% of surface elements specified by each microblogging group were repeated by other groups (compared to 83% in picnics and 80% in wikis). (Appendix F: Numeric Data on Comparative Evaluation TAPT Artefacts provides a full breakdown of the coded TAPT artefacts.) As shown in the appendix, microblog analyses had few unique literal effects (11, opposed to 22 for picnics and 18 for wikis), suggesting that microblog groups found it harder to come up with effects. This is perhaps a reflection of the phenomenon's youth.

5.3.2 Picnics

The six picnic artefacts designed using TAPT consisted of:

- Two websites to deliver food for consumption while conferencing
- One online picnic shop with a few web games/social interactions
- One online space with various 'locations', realtime chat and a shop (no mention of food)
- One public virtual area: virtual food, optional chat/activities
- One live chat site with picnic-related alerts (bees, rain).

An example TAPT design was:

A website that allows friends to 'meet' in a 2D or 3D virtual space. Environment could be customised by host/users eg. beach, field etc. real time discussion. Ideally would be integrated with other social networks e.g. Facebook to allow inviting and creation/scheduling of the event.

Product can be selected from company catalogue, e.g. picnic rug and ordered if desired. Birdsong/background noises played to participants alongside speech.

Common surface elements were food, friends, outdoors, rugs and baskets. Literal effects often involved being outside or experiencing weather, eating and conversation, while abstract effects involved sharing, companionship and relaxation. Distilled experiences focused upon the shared social experience and food. Implementations were fairly similar, and generally somewhat disappointing: the picnicking task was chosen to be challenging, but responses generally involved enabling people to eat at their computers while conferencing, rather than replicating picnic experiences in other, less obvious ways.

Picnics had the greatest number of literal effects per group (6, compared to 4 for each of the other tasks), and the greatest proportion of key effects to be identified as key by more than one group (73%, as opposed to 33% for wikis or 43% for microblogs). This suggests that participants found identification of literal effects of picnicking easier than identification of such effects when using wikis or microblogging.

Participants mentioned difficulties inherent in the picnic task. For example, in answer to the question on how well their design recreates the original experience, a Scenarios user (Participant 2) said "Not well as a picnic is a very physical experience rather than a virtual so was hard to recreate." TAPT users said similar things: for example, Participant 24 said "The original is all about physical experiences (food + environment) so difficult to reproduce well." The issue of the difficulty of the scenario arose in a group discussion:

Participant A: I couldn't really comprehend how you could translate what we were trying to do into the new medium ... but whether that's a fault of the methodology or that scenario doesn't translate very well or whether it's just me

Participant B [from a different group]: did you have the picnic?

Participant A: yes

Participant B: yeah, I did think that'd be quite hard to do

Participant A: it's a computer assisted picnic, and it has to involve food!

The building of shopping-based websites is of interest: it initially seemed that groups had either misunderstood the task, or just 'given up' and built shops. However, at least some of these designs were built as shops with intent. For example, Group I listed 'shopping' as a literal effect in their analysis, while Participant 36, remarked that the design was better "Because when they buy anything in a real store, it's hard to imagine how to use the stuff they buy." Notably, in a later expert review one expert thought that incorporating business and money-making aspects was part of the task (see Section 6.5.1).

5.3.3 Wikis

Of the 13 wiki artefacts, one TAPT design and one Unstructured Discussion design were incomprehensible (the second item in the list of TAPT-generated wiki designs, and the second item on the list of Unstructured Discussion-generated wiki designs in Appendix G: Artefacts from the Comparative Evaluation. These two items are excluded from the analysis.

TAPT-generated wiki designs were relatively similar. Of the five clear TAPT designs, two used interactive walls, one a non-interactive wall, one a touchscreen, and one an interactive whiteboard. One example design was:

> *An interactive white board that is editable by anyone who picks up the pen. Content is archived and old entries fade to create spaces to entice new contributions. Users may choose to translate others' experiences or add photos. Content is sent to a traditional wiki for sharing and collaborating remotely.*

Initial wiki experience descriptions mentioned community editing, content being accessible to all, and reviewing of content. Common literal effects centred upon acquiring information, while abstract effects focused upon satisfaction, shared experiences / collaboration, and increased knowledge / learning. Distilled experiences concerned shared collaborative aspects.

Wiki groups chose fewer effects as key (19%, compared to 43% for microblogs and 47% for picnics), and had no overlap of key effects (compared to 33% for microblogs and 40% for picnics). This implies difficulty in agreeing on what is a key literal effect of wiki use.

Some participants found the wiki task challenging, as described in a group discussion:

> *One thing I felt was that … with the care homes one […] it was quite easy to sort of provide a natural analogous experience in the care home whereas we go on to do the Berlin wall thing wiki … whereas for that environment it was harder to think of something which was … similar to the way a wiki works […] unless you were providing a screen with a wiki on it.*

Given the prevalence of the “screen with a wiki on it” design format, it seems not unreasonable to assume that this participant was voicing an opinion shared by others.

TAPT users seemed particularly driven by wiki-like concepts. For example, 4 out of 5 TAPT designs (and also 2 out of 4 Unstructured Discussion designs) made it possible for museum visitors to edit one another’s shared experiences. Although at times editing is appropriate (for example, staff removing offensive content), allowing visitors to edit or delete one another’s content seems inappropriate for this scenario, which is about allowing visitors to comment and share their experiences.

One participant who used Unstructured Discussion, when commenting on replication of social aspects, remarked “User expectation will be more in line with a guest book or journal which will allow them to leave a permanent record of their thoughts”. This leads to an important point about the wiki scenario: a collaborative area for sharing and commenting on subjective experiences is not the same as an editable wiki designed for sharing facts. Accordingly, it seems likely that a successful implementation of this scenario would not necessarily involve blindly copying surface elements of a wiki. To further explore this, Figure 5-2 shows wiki participant responses to the question: ‘How well did you replicate the surface design elements of the original design experience?’

As can be seen, the participants using Scenarios tended not to answer with very strong feelings, although most felt they had done ‘well’ at replicating the design elements of a wiki. More interesting are the results from participants using TAPT and Unstructured Discussion, who seemed evenly split on how ‘well’ they did at this task.

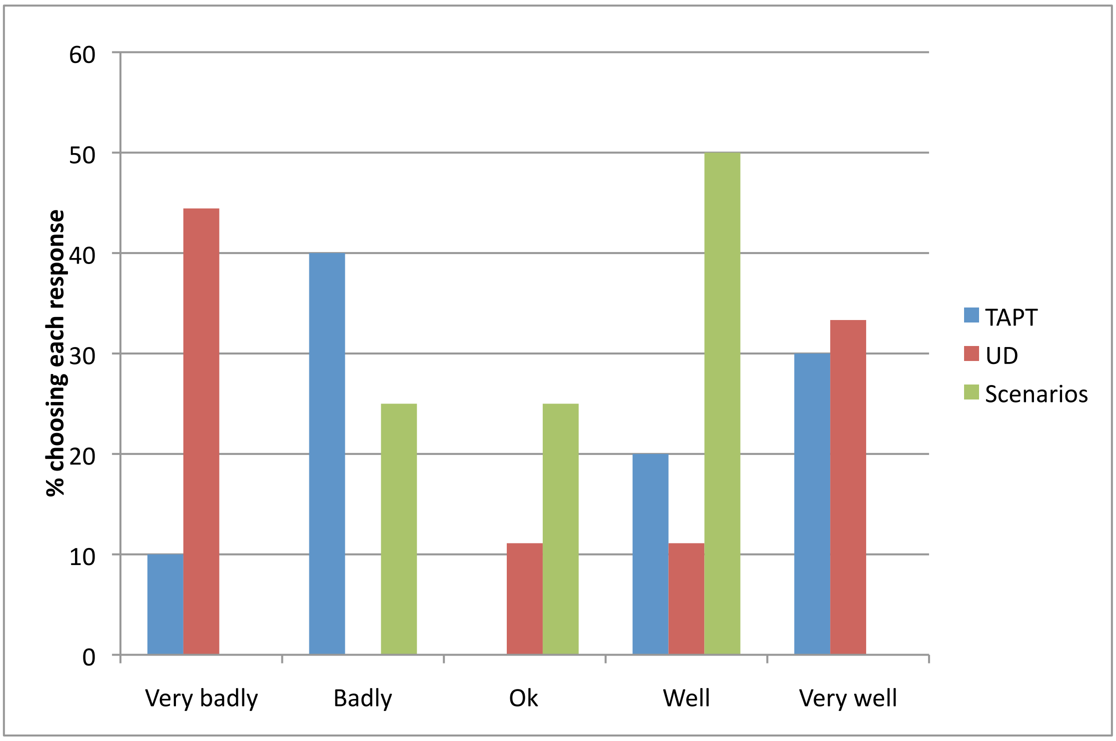


Figure 5-2 Wiki participants' responses on replication of surface elements

The question on replicating surface design elements perhaps left participants feeling they 'ought' to answer positively ('well' or 'very well'). During the trial, at least one participant asked the investigator whether it was 'ok' to tick 'very badly' in answer to this question when they were happy with their design. This possibility of varying interpretations of the question is further supported by some of the comments: one participant who ticked 'very badly' and used Unstructured Discussion left a positive remark of 'Stripped out almost all features of a wiki for simplicity', while another who also ticked 'very badly' and used Unstructured Discussion said, negatively, "I can't see the relationship between the redesign and the original experience".

When carrying out the wiki task, some TAPT participants found Teasing Apart difficult: comments included "I found it difficult to put into words some of the aspects of a generic wiki environment" and "Surface elements & literal effects are difficult to separate." Others were positive: "Breaking the system down into surface elements and emotional side-effects helped us to know how people used a wiki and why"; "The problem was broken down into smaller seemingly unnecessary steps which actually helped in understanding and approaching the problem at hand."

Participants particularly expressed difficulty considering the emotional effects of wikis: “Describing the emotional effects of a wiki was hard” and “A wiki seemed a very sterile experience to assign any emotional effects to”. Although some people found this aspect difficult, considering emotional aspects is a necessary step in rebuilding experiences. Positive comments about the prompt included: “Had not specifically considered emotional aspects before” and “Did make me think of [emotional effects] where I usually wouldn’t”

In summary, the wiki task led to various designs that were often but not always wall-based. Participants appeared to struggle with the transition from a web-based, factual wiki to a physical area for sharing subjective materials.

5.4 Analysis of Participant Perceptions

This section discusses participant responses to the questionnaire and their comments in post-study group discussions. The section considers in turn comments (and, where they were given, ratings) on: understanding experience; replicating experience; structure and scaffolding; creativity and diversity of designs; workplace suitability; responses to the comparative question; accessibility; the impact of learning time and the time required to apply methods.

This section presents paired t-tests of participant responses to questions, comparing responses to the three methods which participants used.

Due to the quantity of data gained, it is not practical to present all qualitative data. Appendix H: Sample Questionnaire Response presents a sample questionnaire response, while Appendix I: Comparative Evaluation Ratings summarises ratings given by participants.

5.4.1 Understanding experience

Table 5-14 summarises participants’ ratings of methods’ usefulness for understanding experience, overall and in terms of emotional aspects, social context, and user perceptions and expectations, where a score of 1 equates to ‘very bad’ and a score of 5 equates to ‘very good.’

	TAPT	Scenarios	UD
Overall	4.03	3.46	3.41
Emotional	3.78	3.25	2.93
Social	3.61	3.39	3.15
Perceptions and expectations	3.23	3.21	2.44

Table 5-14 Responses on understanding experience

It is interesting that the overall average score (across all four rows) is higher than the marks for emotional, social and process-related aspects. This implies that either the methods somehow afford an overall understanding greater than the sum of their parts, or the questionnaires did not ask about an important aspect.

Of particular note is that participants found TAPT the most useful method overall. This result has statistical significance, shown in Table 5-15.

	TAPT	Scenarios	UD
Mean	4.03	3.46	3.41
Standard deviation	0.82	1.02	1.08
Variance	0.68	1.02	1.17

Table 5-15 Statistics on methods' aid to understanding

TAPT > Scenarios ($t=2.21$, $df=54$, $p=0.025$); TAPT > UD ($t=2.41$, $df=57$, $p=0.01$);

Scenarios > UD ($t=0.17$, $df=49$, not significant)

Comments focused on TAPT's analytical qualities ("Separating literal & abstract was useful to improve understanding"; "It made me analyse what was happening"; "Helped understand key aspects"). This is in contrast to Scenarios ("Didn't really cover the original experience. Focused on the new"; "Needed analysis to think of personas") and Unstructured Discussion ("It was very easy to skip past the original experience"; "We didn't analyse [the original experience]"). It seems that participants felt that TAPT's analysis phase enabled deeper understanding of the original experience.

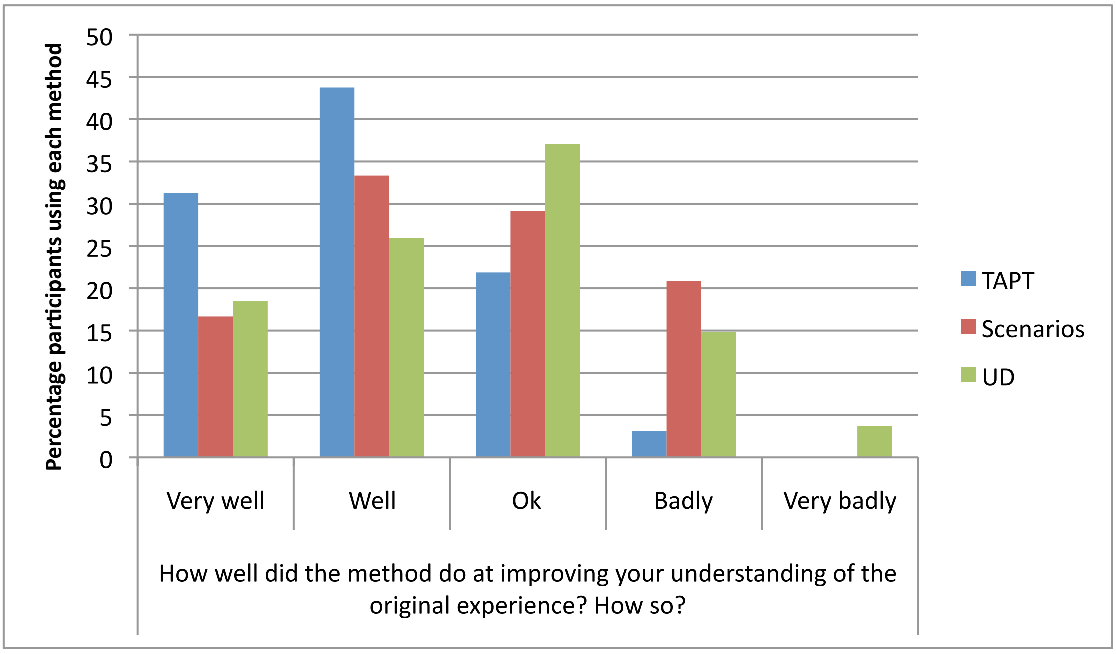


Figure 5-3 Responses on understanding experience

The next three questions (on emotional, social and expectation-related aspects) concerned a subset of overall usefulness, and responses followed a similar pattern.

In terms of emotional aspects, TAPT was mostly highly rated, and again this rating was statistically significant (see Table 5-16). Comments suggest that the high rating is due to TAPT's analysis phase ("It made you think about what emotions we experienced when using a wiki that you don't realise you experience."; "This has opened my eyes to see why these feelings come into play").

	TAPT	Scenarios	UD
Mean	3.78	3.25	2.93
Standard deviation	1.04	1.03	1.30
Variance	1.08	1.07	1.69

Table 5-16 Statistics on awareness of emotional effects

TAPT > Scenarios ($t=1.87$, $df=54$, $p=0.05$); TAPT > UD ($t=2.71$, $df=57$, $p=0.005$); Scenarios > UD ($t=0.97$, $df=59$, not significant)

Scenarios users were less enthused ("Doesn't encourage thought about the original experience"), although were somewhat positive ("Get to see why people actually use it and how it affects them."). Users of Unstructured Discussion were generally negative about the lack of prompts to consider this area ("Didn't spend much time talking about this"; "We assumed we were v. familiar with the original experience & didn't analyse it.").

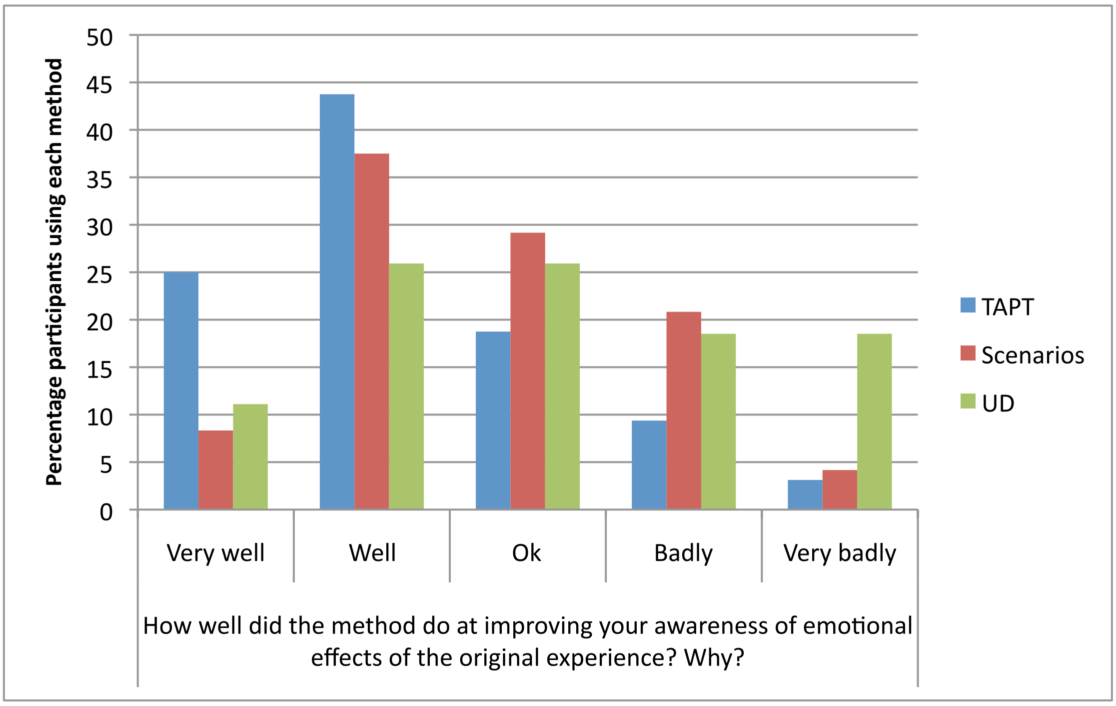


Figure 5-4 Responses on awareness of emotional effects

Concerning understanding social context, the three methods were more closely ranked, although again TAPT led. It is likely that Scenarios did better here due to its user focus. When applying an alpha level of 5% or less, there was no statistical significance to TAPT's lead (Table 5-17).

	TAPT	**Scenarios**	**UD**
Mean	3.61	3.39	3.15
Standard deviation	1.09	0.89	1.12
Variance	1.18	0.79	1.26

Table 5-17 Statistics on awareness of social effects

(TAPT > UD t=1.53,df=55; TAPT > Scenarios t=0.81,df=52;

Scenarios > UD t=0.81, df=46, not significant)

However, TAPT and Scenarios were found to be statistically significantly better than Unstructured Discussion (to an alpha level of 0.5%, and 0.25% respectively) for a certain scenario. This was when considering the half of responses from people who used the methods second in the study, shown in Table 5-18.

	TAPT	Scenarios	UD
Mean	3.76	3.50	2.67
Standard deviation	0.97	0.85	0.78
Variance	0.94	0.72	0.72

Table 5-18 Statistics on awareness of social effects, second task:
TAPT > Scenarios (t=2.71, df=25); TAPT > UD (t=3.25, df=27, p=0.005);
Scenarios > UD (t=2.27, df=20, p=0.025)

One could speculate from this stark difference that perhaps users of Unstructured Discussion judge the approach more harshly when they are using it after having used a more structured method. That is, if Unstructured Discussion is the first method used, participants approve of it more strongly, but when a more structured method has been used beforehand, participants find Unstructured Discussion more disappointing. Whatever the cause of the difference, Table 5-19 confirms that the set of participants to use Unstructured Discussion second definitely rated it less well than those who used it first.

	Very well	Well	Ok	Badly	Very badly
First method used	5	4	5	1	0
Second method used	0	3	5	3	1

Table 5-19 Ratings of Unstructured Discussion, showing whether it was used first or second

Positive comments on TAPT's use in the context of social understanding again related to analysis: "Drew out essentials"; "Formed a lot deeper meaning to social side of picnic"; "Got us to think things through". Some negative comments emerged: "Didn't consider goals, motivation etc" and "Didn't focus on social context". Scenarios users were also split: one said "didn't really cover the original experience" compared to "This method is geared directly at doing this". Those who used Unstructured Discussion were more negative: "We failed to consider this factor".

Regarding understanding perceptions and expectations, TAPT and Scenarios were rated similarly, ahead of Unstructured Discussion: probably this was due to the advantages of Scenarios' user-focus and TAPT's analysis. There was no statistically significant difference between TAPT and Scenarios in this context, although Table 5-20 demonstrates that TAPT's ranking was significantly better than Unstructured Discussion's rating.

	TAPT	Scenarios	UD
Mean	3.23	3.21	2.44
Standard deviation	1.01	1.02	1.02
Variance	1.01	1.04	1.04

Table 5-20 Statistics on awareness of perceptions and expectations

TAPT > UD ($t=3.15$, $df=55$, $p=0.005$); TAPT > Scenarios ($t=0.09$, $df=52$, not significant); Scenarios > UD ($t=2.83$, $df=49$, $p=0.005$)

Positive comments on TAPT centred on increased knowledge ("Made you think about using a wiki differently"; "More thought put into how OAP users would view this phenomenon"), although some commented that TAPT could focus more on changes in user state or on users in general. Scenarios users commented positively on its user-focus, but some felt it did not necessarily help or cover this area. One Unstructured Discussion user "thought extensively about user attitudes, purposes", but in general comments on this method were negative, presumably due to the lack of prompts to consider this area.

Overall comments on TAPT included remarks that "Teasing Apart was a great structured way of getting to the essence of an experience" and "Teasing Apart is useful for understanding the core aspects of what makes a system usable and what people want."

In summary:

- TAPT was rated as better than Scenarios and UD at improving overall understanding (with statistical significance)
- TAPT was rated as better than Scenarios and UD at improving understanding of emotional aspects of experiences (with statistical significance)
- TAPT was not rated significantly better than Scenarios or UD at improving the understanding of social aspects, but TAPT and Scenarios were statistically better than UD in the context of participants conducting the second task
- TAPT was statistically better than UD in terms of understanding expectations of experiences.

Users consistently gave UD lower marks after using a structured method (TAPT or Scenarios) indicating that they did not realise before using them how the structured methods might be of benefit.

5.4.2 Replicating experience

Table 5-21 summarises responses on experience replication.

	TAPT	Scenarios	UD
Overall	3.61	3.63	3.67
Emotional	3.55	3.50	3.31
Social	3.62	3.57	3.63
Experience as a process	3.47	3.39	3.00
Superficial replication	3.48	3.26	3.19

Table 5-21 Responses on replicating experience

As can be seen, the results were much more closely ranked for replicating experience than for understanding experience. As will be discussed, only one statistically significant result was found in this set of results, which was that TAPT was found to be better than UD for replication of experience as a process.

	TAPT	Scenarios	UD
Mean	3.61	3.63	3.67
Standard deviation	0.99	0.77	0.83
Variance	0.98	0.59	0.69

Table 5-22 Statistics on replicating experience

TAPT > UD (t=0.22, df=56); TAPT > Scenarios (t=0.05, df=53); Scenarios > UD (t=0.18,df=49). No significant results.

Negative remarks from TAPT users mostly concerned the nature of the original experience (e.g. “The wiki concept does not translate well to a physical museum”; “The original is all about physical experiences (food + environment) so difficult to reproduce well”), while positive comments concerned successful re-provision of key features of the original experience. Most Scenarios users were positive (“I think it does cover all of the original experiences”), and Unstructured Discussion users were very confident, with comments such as “We think it does [replicate the original experience]”; “Nearly all aspects covered”; “It’s about as good as you could expect from a website”.

TAPT’s low ranking here may seem surprising as it is intended to capture experience. This result may be because it encourages people to *reimagine* or translate an experience, making it superficially less like the original. One could speculate that after carrying out a deep analysis of the starting experience, TAPT users were particularly aware of aspects which they had *not* replicated, and therefore rated their results less well than otherwise. This is indicated by participant ratings of TAPT, who did not rate it as the best method at replicating experiences overall, but rated it highly in terms of replicating specific aspects (emotional, social, the experience as a process). This suggests they understood the broad question "Did this process help you replicate the experience?" in a very literal sense, and thus rated TAPT less well.

Likewise, the confidence exhibited by Unstructured Discussion users may be borne of the lack of structured thinking provided by the other two methods: having not analysed the starting experience, Unstructured Discussion users were less aware of deeper experiential aspects which may not have been included in their designs.

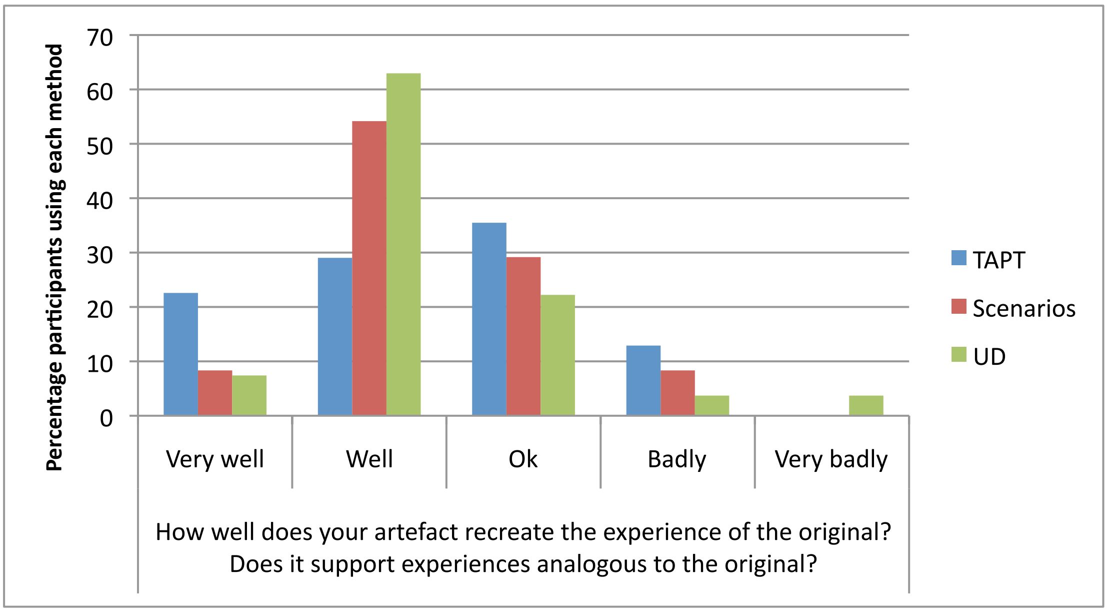


Figure 5-5 Responses on recreating experience

Concerning replication of emotional aspects, all three methods were similarly rated. Notably, Unstructured Discussion had the most ratings of 'Very Well', perhaps reflecting the greater confidence of users of Unstructured Discussion. There were no statistically significant findings, shown in Table 5-23.

	TAPT	Scenarios	UD
Mean	3.55	3.50	3.31
Standard deviation	0.77	0.88	1.01
Variance	0.59	0.78	1.02

Table 5-23 Statistics on replicating emotional aspects

TAPT > UD ($t=0.98$, $df=55$); TAPT > Scenarios ($t=0.21$, $df=53$); Scenarios > UD ($t=0.70$, $df=48$). No significant results.

In terms of replicating social context, as with understanding social context, all three methods were rated similarly: this could be because the three tasks all concerned social situations, and thus perhaps the subject matter prompted consideration of this area regardless of method. There were no statistically significant findings.

	TAPT	Scenarios	UD
Mean	3.62	3.57	3.63
Standard deviation	0.86	0.95	1.04
Variance	0.74	0.89	1.09

Table 5-24 Statistics on replicating social aspects

TAPT > UD (t=0.03, df=54); TAPT > Scenarios (t=0.21, df=50); Scenarios > UD (t=0.22, df=48). No significant results.

Concerning replication of changing expectations and perceptions, the methods were again similarly rated, although TAPT and Scenarios were never rated as doing 'Very Badly' at this. TAPT was statistically significantly better at this facet than Unstructured Discussion.

	TAPT	Scenarios	UD
Mean	3.47	3.39	3.00
Standard deviation	0.73	0.79	0.89
Variance	0.53	0.62	0.80

Table 5-25 Statistics on replication of expectations and perceptions

TAPT > Scenarios (t=0.35, df=51, not significant); TAPT > UD (t=2.09, df=42, p=0.025); Scenarios > UD (t=1.59, df=47, not significant)

Finally, participants were asked how well they felt their designs did at recreating superficial aspects of the starting experiences. Table 5-26 shows responses.

	TAPT	Scenarios	UD
Mean	3.48	3.26	3.19
Standard deviation	1.09	1.01	1.41
Variance	1.19	1.02	1.99

Table 5-26 Statistics on superficial replication

TAPT > UD (t=0.83, df=53); TAPT > Scenarios (t=0.74, df=50); Scenarios > UD (t=0.19, df=47). No significant results.

Many TAPT users rated the method as doing well at replicating surface design elements (although the result was not statistically significant). The researcher did not expect TAPT to do well at this: users appeared to dislike rating TAPT as 'bad' in this respect as they chose not to replicate surface elements. Saying it was 'bad' or 'very bad' implied that they wanted to achieve this goal but could not, rather than choosing not to (indicated by comments such as "We deliberately tried to be different from the wiki" and "We intentionally changed them all").

In summary, participants rated:

- UD as best at overall replication (with no statistical significance)
- TAPT as best at replicating emotional aspects (with no statistical significance)
- UD as best at replicating social aspects (with no statistical significance)
- TAPT better than UD for replicating changing expectations and perceptions (statistically significant)
- TAPT as best at replicating superficial aspects (with no statistical significance)

TAPT users rated it as less good at replication overall, although it received better ratings with specific facets, implying that TAPT was perceived as helpful at replicating particular underlying aspects of experiences.

5.4.3 Structure and scaffolding

Table 5-27 shows participants ratings of methods' efficacy at scaffolding discussions. TAPT and Scenarios were both statistically significantly better at this than UD.

	TAPT	**Scenarios**	**UD**
Mean	4.25	4.04	2.93
Standard deviation	0.76	0.86	0.86
Variance	0.58	0.74	1.84

Table 5-27 Statistics on scaffolding discussions:

TAPT > Scenarios (t=0.92, df=54, not significant); TAPT > UD (t=4.43, df=57, p=0.0005);

Scenarios > UD (t=3.48,df=49, p=0.001)

Some TAPT participants reported finding the structure helpful in the short timeframe: "Formal structure was helpful in short timeframe" (Participant 3); "Structure made it easier to get started" (Participant 11); "Following a pre-defined structure always helps to tackle problems more effectively and in the shortest possible time" (Participant 15). A Scenarios user found the same thing ("It was good but hard to know how much time to spend on each section", Participant 25).

Some TAPT-users commented on a disconnect between the output of Teasing Apart and the input to Piecing Together. For example, "Piecing together was hard, felt hamstrung by the distilled experience we had come up with and struggled to replicate it." In the group discussion, another comment was:

> *We kind of struggled with TAPT [...] I really liked the Tearing Apart [sic] phase, we did a very good job of like breaking down the Wikipedia experience*

into sort of very ... granular specific level, but then we ended up we realising as we came to do the Piecing Together that ... a lot of the very low level specific elements didn't really translate very well [...] we realised that we needed to back out and take a more general approach ... to solve the problem and we'd spent a lot of time building up these little specific blocks

These responses alongside observations from the expert review about a disconnect between Teasing Apart analyses and Pieced Together designs (see Section 6.4.3) suggest that perhaps some work is required to aid participants in crossing this gap.

Some designs from groups using Unstructured Discussion showed signs of structured analysis. This includes three of the four Unstructured Discussion groups responding to the microblogging task: group L included headings (‘assumptions’ and ‘design’); group Q included some explicit reasoning (‘very little “social networking”, to be done, as the social network is already enforced by the environment. The system therefore needs to act as…’); group T included assumptions (‘We assume that…’). Group U, who applied Unstructured Discussion to wikis, also included such text. These responses suggest that although participants were not allowed to write or sketch ideas, at least some groups’ discussions had a level of structure.

Participants provided both positive and negative comments about Unstructured Discussion. Some participants preferred open discussion because it “facilitates quick thinking and creativity”, and liked being able to “think very openly”. However, others noted the risk of missing important areas (“easy to get sidetracked”) and the ‘randomness’ resulting from the lack of structure (“Too freeform”, “We chose discussion points at random”). One participant summarised, saying: “structure => thoroughness, no structure => creativity”. It would appear that participants using Unstructured Discussion tended to find that it enabled creativity and ‘wild ideas’, but found a flimsiness in its lack of structure, and omitted important aspects too easily.

In summary, participants found TAPT and Scenarios better than UD for scaffolding discussion, with statistical significance.

5.4.4 Creativity and diversity of designs

As shown in Table 5-28, participants rated Unstructured Discussion as the best at supporting creativity, although this result was not statistically significant. Comments from Unstructured Discussion users suggested that they found its lack of structure helpful for supporting creativity (addressed in Section 5.4.3).

	TAPT	Scenarios	UD
Mean	3.77	3.79	4.15
Standard deviation	0.92	0.93	0.95
Variance	0.85	0.87	0.90

Table 5-28 Statistics on creativity

TAPT > UD (t=1.49, df=56); TAPT > Scenarios (t=0.07, df=53); Scenarios > UD (t=1.33, df=49). No significant results.

Diversity of resultant designs is another way to gain insight into the creativity fostered by the three methods, as more creative approaches presumably result in more diverse designs.

Designs varied in two senses: the mode of implementation and the functionality provided. For example, many wiki designs were wall-based (mode of implementation), while many picnic designs varied between focusing on enabling online conferencing and online shopping (functionality). Variation in functionality is of less interest, as this happens when groups respond to a different perceived task, but variation in modality is relevant. As the design task for picnicking requested a website-based implementation, picnic responses are not of interest, but one can consider microblog and wiki responses.

Table 5-29 shows the variety of designs for the microblog and wiki task. It lists the number of different concepts (such as a voice system as opposed to a tablet-based system) and the number of resultant designs.

	TAPT	Scenarios	UD
Microblog	4/5 (80%)	3/5 (60%)	3/4 (75%)
Wiki	2/6 (33%)	2/2 (100%)	2/4 (50%)

Table 5-29 Diversity of microblog and wiki designs

(number different concepts / number designs)

Microblog designs were varied across all methods. TAPT wiki responses were rather similar (two different concepts emerging from six different designs), while Unstructured Discussion wiki designs were somewhat more varied, and the two

Scenarios wiki designs were both different. With so few wiki designs generated by Scenarios, it is difficult to comment on this result with confidence, but TAPT and Unstructured Discussion certainly resulted in many wall-based designs: 5/6 TAPT designs used walls, and 3/4 Unstructured Discussion designs did the same. This homogeneity across both TAPT and Unstructured Discussion suggests that the wiki scenario perhaps led participants to build wall-based designs: evidence from the expert review also suggests this. See Section 6.5.1.

In summary, participants rated UD as best at supporting creativity, although this result did not have statistical significance. Designs produced with each method varied to some extent.

5.4.5 Workplace suitability

Participants felt that output from TAPT and Scenarios was much more suited to workplace use than output from Unstructured Discussion and their ratings of the more structured methods were statistically significantly better than those of Unstructured Discussion (Table 5-30). Comments on TAPT noted its intuitive layout and the ability to trace decisions. Some participants suggested allowing a space for flow charts and diagrams.

	TAPT	**Scenarios**	**UD**
Mean	3.50	3.58	2.65
Standard deviation	0.88	0.72	1.02
Variance	0.77	0.51	1.04

Table 5-30 Statistics on workplace suitability: TAPT > UD ($t=3.28$, $df=56$, $p=0.001$); TAPT > Scenario ($t=0.38$, $df=54$, not significant); Scenarios > UD ($t=3.68$, $df=48$, $p=0.0005$)

In summary, participants found TAPT and Scenarios more suited to workplace use than UD, with statistical significance.

5.4.6 Comparative question

At the close of the study, participants were asked which of the two methods they had used they found easier to apply, which was more useful, and which was better for replicating experiences. Some participants felt unable to answer, because the tasks were so different: “The scenario (social networking) itself was more amenable to replication than the wiki one"; "I think the second task hindered TAPT's potential strengths"; "The experience was almost impossible to replicate fully."

Table 5-31 shows, for each pair of methods, the number of people to rate each of the compared methods as better.

	Scenarios and UD		**Scenarios and TAPT**		**UD and TAPT**	
	S	UD	S	T	UD	T
Ease of use	4	5	7	7	12	4
Useful results	7	1	4	8	4	9
Replication	5	5	1	11	5	7

Table 5-31 Comparative question

Given that Unstructured Discussion consisted of talking about ideas before writing down a design (in contrast to the structured, step-by-step approaches of Scenarios and TAPT), it is unsurprising that participants found Unstructured Discussion easiest. Scenarios was found easier to apply than TAPT, probably because many participants were already familiar with the technique.

TAPT was ranked as the most useful method, followed by Scenarios, followed by Unstructured Discussion.

Participants found TAPT the most useful for replicating experiences, with Scenarios rated better than Unstructured Discussion. TAPT's high rating at replication is surprising when considered in conjunction with its poor ratings in the task questionnaire (Section 5.4.2). This difference is likely due to changing perceptions of participants over the course of the study.

5.4.7 Accessibility

Scenarios is a user-focused approach, which one might expect to result in more accessible designs.

Microblog designs generally seemed more accessible, perhaps because the task itself more explicitly highlighted issues of accessibility: when end users live in a care home, it is straightforward to reason that disabilities and accessibility are likely to be an issue, whereas when catering to museum visitors or picnickers, this may be a less obvious facet. Nonetheless, many microblog designs appeared to cater poorly for impaired users.

One measure of accessibility is to consider how many designs allowed input via more than one modality. Where descriptions were unclear, they were considered not to offer multimodal input. Examples include Group I's microblog design ("Interactive TV, menu-based interface to enable phone calls, viewing of news/images, discussions (input method unclear)") and Group L's picnic design ("Friends 'meet' online in a 2D or 3D space, with customised environment such as a beach or field. Real-time discussion occurs, and users can buy products. May be integrated with Facebook. Background noises like birdsong included").

Table 5-29 shows the results of this count.

	TAPT	Scenarios	UD
Microblog	2/5 (40%)	3/5 (60%)	1/4 (25%)
Picnic	2/6 (16%)	2/5 (40%)	4/4 (100%)
Wiki	2/6 (33%)	2/2 (100%)	2/5 (40%)

Table 5-32 Multimodality of designs (number multimodal / total)

In the picnic task, Unstructured Discussion resulted in the most multimodal results: this is probably related to the functionality produced, as all four Unstructured Discussion artefacts consisted of a website for conferencing while eating, while TAPT and Scenarios artefacts included online shops and planning tools. Accessibility was rarely a focus of picnic designs, even those generated with Scenarios.

Scenarios generated the greatest proportion of multimodal microblog and wiki artefacts, with TAPT and Unstructured Discussion performing approximately equally across the two tasks. Despite this, many of the Scenarios-generated microblog designs only really accommodated people able to use relatively complex systems: designs 3 – 5 in the Microblog Scenarios list of Appendix G: Artefacts from the Comparative Evaluation all appear to involve keyboards or complex menu structures. This is not necessarily an unhelpful approach, but some of the details implied complicated interfaces which may be unsuitable for a frail and elderly audience, for example requiring keyboard input or the use of animations. This is despite Scenarios groups building personas with various disabilities: perhaps an indication that designers need a greater awareness of the impact of these disabilities.

A key aspect of the use of personas is that it explicitly prompts designers to consider the various types of user who will use the end system, and to cater for their needs – as reflected in participant feedback. Comments included that "creating the personas was quite hard" because "it was difficult to pinpoint what a person would be like."

One way to improve the Scenarios approach would be to explicitly prompt designers to consider different groups, for example by including a prompt such as

"Don't forget to consider different groups, such as older or younger users, users with physical or mental disabilities, or secondary stakeholders such as staff or family."

In summary, the accessibility of designs appears to vary across both tasks and methods. Microblog designs seemed particularly accessible, perhaps due to the nature of the task, and Scenarios resulted in the highest proportion of accessible microblog designs, followed by TAPT.

5.4.8 Time required to learn and apply methods

Participants felt 'pushed for time' with the two structured methods, particularly TAPT. For example, when answering the question "Were you able to carry out the task?", five TAPT-users remarked on the shortness of time with comments such as: "Ran out of time in Piecing Together" (Participant 8); "Could have used much more time" (Participant 9); "Teasing apart took a long time" (Participant 20). Some Scenarios users made similar comments, such as Participant 42: "it was hard […] not [to] spend too much time on personas even after they'd served their purpose."

By contrast, participants found UD much faster to apply: "Discussion was much quicker" (Participant 33); "We didn't spend much time to start" (Participant 36); "More time to investigate ideas" (Participant 20).

Finally, some participants commented that they felt short of time with TAPT specifically because of the learning curve: ("Learning terms […] takes time", Participant 16); "Not enough time (only 30 – 40 minutes including reading and understanding material)" (Participant 31); "I think once the TAPT procedure was second-nature, it could be used […] very well" (Participant 33).

Participants were generally familiar with the concept of Scenarios, but with the exception of Group I, who applied TAPT twice (and whose responses are excluded here), all TAPT participants were using the method for the first time. It could be that introducing a 'trial run' for familiarisation purposes might have helped these users. Section 6.6 returns to this.

In summary, it appears that UD users were happy with the amount of time they had to apply the method, which is a contrast with users of the other two methods. TAPT users felt particularly pushed for time, probably because it was an entirely unfamiliar method.

5.5 Conclusions

This chapter presented an evaluation of TAPT, a method for understanding and redesigning experiences across different contexts. The results of the trial suggest that the strength of TAPT is its ability to provide a structured method for thoroughly exploring all the factors of a given experience, including those emotional and social effects that might otherwise be overlooked; its primary weakness is that it is not user-focused, which can lead to less accessible designs.

Table 5-33 shows the strengths and weaknesses of the three methods as elicited from participant responses.

Clearly, no one method will suit all circumstances, but knowing the strengths and weaknesses of methods allows us to apply them together in an appropriate way. Some participants alluded to this: “Need a combination. TAPT good for key features. Scenarios good for user relevance.”, “[Unstructured Discussion] is really only one stage in a process.”

It appears TAPT meets the goal of enabling translation of experiences, as abstracting ideas frees users to be creative. For example, one comment was “The distilled experience forced a more abstracted idea, which helped reshape the experience much more easily.” Participants did not always immediately feel comfortable using TAPT, but generally found it useful: "The problem was broken down into smaller seemingly unnecessary steps which actually helped in understanding and approaching the problem at hand."

Participant feedback suggested greater benefit could be gained by applying TAPT at multiple points through the sequence of an experience, and to multiple users within an experience. This returns to the concept of combining methods: for example, one might apply Scenarios and Personas to elicit user groups, and then apply TAPT to each usage scenario.

One participant commented “personal experience makes a huge difference”: it is difficult to analyse experiences with which you are unfamiliar. Another aspect of future work is to consider how groups can apply TAPT, and ideal group size and composition.

Based on these results, possible improvements to the TAPT process include:

- providing lists of possible experienced effects, as a prompt
- encouraging reflection upon which experiential aspects designers may wish to omit
- allowing practitioners to include diagrams and sketches in the Piecing Together phase

Method	Strengths	Weaknesses
TAPT	• Builds understanding (5.4.1, 5.4.6) • Supports replication of underlying experiences (5.4.2, 5.4.6) • Structured, systematic, logical (5.4.3) • Can foster creativity (5.4.4) • Suitable for workplace use (5.4.5) • Documentation of decisions (5.4.5)	• Structure can distract from creativity (5.4.3) • Lacks user focus (5.4.7) • Learning curve (5.4.8) • Takes some time to apply (5.4.8)
Scenarios	• User focus (5.4.7) • Structured, systematic (5.4.5) • Can foster creativity (5.4.4) • Suitable for workplace use (5.4.5)	• Lacks focus on the experience (5.4.1) • Structure can distract from creativity (5.4.3) • Takes some time to apply (5.4.8)
UD	• Fosters creativity (5.4.4) • Easy to use (5.4.6) • No process to distract (5.4.4), (5.4.8)	• Lacks focus on the experience (5.4.1) • Lacks user focus (5.4.7) • May lead to over-confidence in results (5.4.2) • No structure or focus (5.4.3) • Output unsuitable for workplace use (5.4.5)

Table 5-33 Strengths and weaknesses of methods

Also of interest would be an investigation of the way in which practitioners apply the Teasing Apart phase when they do not know the end domain into which the experience will be Pieced Together.

TAPT was used to understand and re-design various physical and digital experiences: it appears to be an effective tool for analysis and design that could complement existing methods. Teasing Apart physical experiences allows us to implement richer, accessible web-based interactions inspired by these experiences, and TAPT is a valuable tool for re-providing experiences such as browsing social websites via novel, accessible mechanisms.

The method used in this chapter has two weaknesses: participants' assessments of their own designs were likely to be subject to bias, and the experiment itself was run in an artificial laboratory context. Chapter 6 presents an expert review of the output of this experiment in an attempt to gain more objective insight into the resultant designs, while Chapter 7 describes four case studies of TAPT's use in the field.

Chapter 6 Expert Review

Chapter 5 described a comparative evaluation of TAPT against two other design approaches, Scenarios and Unstructured Discussion (UD). Outputs of the evaluation included 42 design artefacts generated by all three methods and questionnaire responses (comments and ratings) on methods' usefulness and ability to support experience replication.

During the course of the study, experts' opinions of these artefacts may have been coloured by the methodology used, other group members, and the subjectivity which results when assessing one's own work. To counter this, the author conducted a separate evaluation of design artefacts in a blind expert review, to more objectively establish their quality and translation of experiences, and to elicit further information about facets such as how creative and accessible artefacts are. In particular it was hoped that this would indicate whether UD users were over-confident about how well their artefacts replicated experience.

This chapter is divided into six sections. First, the experimental design is presented, before a discussion of the experts and factors that could impact the results. Results are given in three parts: blind artefacts are discussed in Section 6.3, TAPT artefacts in Section 6.4, and broader insights in Section 6.5.

Some of the results presented in this chapter have been published (Hooper, 2010b) (Hooper, 2010d).

6.1 Experimental Design

This section describes the experimental design, which was a mixed methods approach that utilised qualitative and quantitative data elicited from the six experts. It covers the broad methodology and the process of normalising the design artefacts to be reviewed, so as to obscure the method by which they were produced. It also describes the interview design and the selection of artefacts for review, before reflecting on the qualitative and quantitative processes used for analysing the results and discussing the use of mixed methods for this kind of problem.

6.1.1 Method

The method involved individually showing six experts six blind artefacts produced in their domains of expertise and eliciting their opinions of these artefacts. Each set of artefacts contained two TAPT artefacts, edited to mask their origin. After the initial review, each expert was shown the TAPT artefacts in full for further assessment.

Resultant design artefacts fell into one of three domains: microblogging, picnicking or wikis. Two experts for each domain were sought, to provide multiple perspectives on the data.

In the case of microblog design and wiki design, experts in human factors and HCI were sought. Experts with a background in sociology or psychology were sought for the remaining category (picnicking, a socially-oriented experience). Recruitment was conducted by the researcher, who approached suitable individuals to ask if they would be willing to participate. No inducement was used.

Individuals were deemed suitable to participate based on their professional qualifications, and being of a background that meant they were unlikely to be intimidated by the interviewer. Each microblog and wiki expert either possessed or was studying towards a PhD in Computer Science and specialised in a relevant domain within that field (for example, user experience of wikis or the impact of blogging). The experts chosen for the picnic task had at least a Bachelors degree in Sociology or Psychology. The experts were from various contexts, with the majority coming from the University of Southampton, but also from IBM Hursley and the University of Edinburgh.

There was some difficulty in selecting experts with qualifications which were relevant to the three tasks of microblogging, wikis and picnicking. Recruitment of Computer Scientists in the first two cases and of a Sociologist and a Psychologist for the final case provided suitable results and insights, however ideally Web Scientists would have been recruited to work with all cases.

The procedure consisted of a one-to-one verbal interview with each expert, subject to audio recording. A semi-structured interview format was chosen as this enabled the researcher to attain consistency of comments across multiple domain experts (as they answered the same broad questions), but also to follow up on interesting comments that were not anticipated in advance. Alternative procedures such as questionnaires were also an option, but would not have afforded this flexibility to delve more deeply. Focus groups were considered, but it was likely that more data could be obtained by talking to experts on an individual basis than in groups: individual interviews enabled the researcher to give full focus to one expert at a time, and eliminated the risk of more forthright individuals 'drowning out' the opinions of others, or individual views being amalgamated to a 'group view'.

One risk in conducting one-to-one interviews is that of confirmation bias, wherein the interviewee gives responses which are overly positive due to either the interviewee wanting to 'please' or 'help' the interviewer, or to the interviewer asking questions in such a way as to encourage such a positive response. To mitigate this risk, several steps were taken:

Firstly, the interviewer took a highly professional approach when interacting with interviewees. This involved being polite and friendly, but maintaining a distance to reinforce that the interview was about gathering professional opinions, not receiving positive feedback or praise.

Secondly, selection of experts was predicated upon experts being in possession of a relevant degree-level qualification and also being people who would not be intimidated by the interview process (i.e. capable of providing negative feedback and not being likely to be over-awed by the interviewer).

Finally, the interviewer made a conscious effort to ask follow-up questions about negative as well as positive comments during interviews: for example, asking why an artefact elicited anger as well as why another artefact was identified as the interviewee's 'favourite'.

During the interviews, experts were asked to comment on the task set to participants of the comparative evaluation, and on a variety of design artefacts. Six normalised (blind) artefacts were shown (two produced with TAPT, two with Scenarios, two with Unstructured Discussion). 'Normalisation' refers to the fact that the artefacts were presented as plain text, with no indication as to the method by which the design was produced. Experts were then shown the full version of the two TAPT artefacts, with the prior analysis as well as the Pieced Together output. The tasks and artefacts were chosen from the experts' domains, so for example an expert in wiki design was shown wiki-inspired artefacts.

An example blind artefact from the set of microblog artefacts is shown below:

> *A communal 'status wall' in the home's common room/lounge.*
>
> *The residents are provided with index cards with their picture, name & room number on. They can then write messages on these index cards & pin them to the wall. These messages can be comments on other people's messages as well as original updates.*
>
> *Moderated by staff who will remove updates when they are too old.*

The number of artefacts used in the study was chosen based on a pilot in which two TAPT artefacts were explored, first as blind, normalised artefacts, then as full TAPT artefacts. The pilot was intended to ascertain how long interviews might take, and lasted 35 minutes including time to reflect on overall patterns and themes. The researcher wished to keep interviews to 60 minutes to avoid being intrusive. These data suggested six blind artefacts and two TAPT artefacts per interview were suitable.

6.1.2 Normalising the artefacts

The goal was to have a set of artefacts of consistent layout and presentation, to mask the method by which each was produced. To achieve this, it was necessary to present the design artefacts as paragraphs of plain text. Artefacts produced with Unstructured Discussion were already in this format, but Scenarios and TAPT artefacts required some work.

In the comparative evaluation, participants using Scenarios were asked to submit a design description, which they were told could optionally refer to their scenarios and personas. Descriptions produced with Scenarios therefore occasionally referred to material external to the descriptions. A similar effect also occurred with Pieced Together TAPT artefacts: occasionally the 'context' field of the Piecing Together field of TAPT contained additional material. When such information was important, it was aggregated into the final description in as light-handed a way as possible, taking care not to distort the content of artefacts.

Most Scenarios users produced an abstract system description, some of which referred to separate scenarios and personas. For example, Group C produced a microblogging scenario which referred to “Majory”, who was separately described. In situations such as these, the researcher added the minimal relevant detail from the separate persona in order to add the relevant context which was lost when the scenarios and personas are removed: in the above example, editing the text from “Majory has just finished…” to "Majory, an 85 yr old widow, has just finished...". Material from scenarios was also incorporated when relevant: for example, Group E (who also addressed the microblogging task) mentioned a tannoy system in their scenarios, but not their system description. This was added into the system description.

Occasionally the ‘context’ field of Piecing Together tables would contain implementation detail rather than the end domain of implementation. For example, Group G’s wiki artefact had a context of “Large touchscreen/smartboard/graffiti wall in a museum. Located in a public space so can be viewed by many people at a time” rather than “Berlin Wall Museum.”

In these instances, the written description was preceded with the line “Context:” followed by the content of this field.

6.1.3 Interview design

Interviews lasted approximately 60 minutes. They opened with the questions described in Table 6-1: first, a title by which to refer to the expert was agreed, and then the expert was asked to comment on the task given to participants in the comparative evaluation. (The tasks are described in Section 5.2.3)

Question	**Desired data**
What professional title would best describe your role in this area?	Agree an appropriate title by which to refer to the expert (for example, ‘senior lecturer working in HCI’).
What would you say are the key challenges in this task?	Ascertain what is felt to be most difficult about this task.
What kind of approach might you use in response to this task?	Insight into approaches currently in use.

Table 6-1 Expert review: questions on expert title and the design task

Next, the expert was shown a sequence of six blind artefacts (two each from TAPT, Scenarios and Unstructured Discussion), and after reading each artefact was asked to comment on it with the following questions:

Question	Desired data
Can you characterise this artefact? For example, is it surprising, insightful, misleading, meaningful, good?	Adjectives about the artefact. This first question was open to elicit opinions on whatever most struck the expert.
Does the designer exhibit any unusual understanding or perception? Were they innovative? Why?	Comments on innovative, creative aspects.
Do you think the designer has responded to the key challenges in the task?	Comments on whether the artefact addressed the most difficult aspects of the task.
Does the artefact translate the experience? (And is that translation of the deeper, underlying experience – for example, of emotional or social aspects of the experience – or is it a translation of more superficial, design elements?)	Comments on whether the artefact replicated the experience (and if so, whether it was a replication of superficial or experiential aspects).
Do the designers seem to have carried out significant analysis in order to produce this artefact? Do you think they were aware of any assumptions they may have made?	Comments on whether a structured, analytical method appears to have been used.
Is the artefact inclusive or accessible?	Comments on accessibility
Have you any other comments on this artefact?	Catch-all.

Table 6-2 Expert review: questions on individual blind artefacts

After all six artefacts had been discussed individually, several questions were asked about the artefacts as a group:

Question	Desired data
Did you notice any patterns or themes running through the artefacts you have seen today? What? Why do you think they are present?	Comments on any commonalities across the artefacts.
Were you surprised by the homogeneity (or lack thereof) across artefacts?	Comments on whether the expert expected to see commonalities.
Do any of the artefacts you reviewed stand out to you as especially inspired or dull? Why?	Comments on any individual artefacts which stand out from the crowd.
Have you any other comments about the artefacts you have seen?	Catch-all.

Table 6-3 Expert review: questions on all blind artefacts

At this stage, the expert had already seen two blind TAPT artefacts, that is, the Pieced Together text only. Now, the expert was shown the two TAPT artefacts in full, with the Teasing Apart analysis as well as the Pieced Together design. Questions were asked about each artefact in turn:

Question	Desired data
Is your perception of the design artefact different now you've seen the analysis which was conducted towards building that description? Why?	Whether the expert is surprised or has other feelings about the artefact now the process by which the design was processed is available.
Given the analysis shown here, would you have come up with a significantly different design artefact? Why?	Whether the expert would use the TAPT analysis differently to the way in which the Software Engineers did.
The first table lists surface elements, experienced effects and the distilled experience. Were you surprised by any of that content? Were any contradictions present?	Whether the expert found aspects of the analysis surprising or odd.
Were you surprised by what was, or was not, chosen as key?	Whether the expert found aspects of the analysis surprising or odd.
Do you have any other comments about this artefact?	Catch-all.

Table 6-4 Expert review: questions on individual TAPT artefacts

Finally, the expert was asked some questions about both of the TAPT artefacts.

Question	Desired data
Did you notice any patterns or themes running through the TAPT artefacts you have seen today? What? Why do you think they are present?	Comments on commonalities across the artefacts.
Were you surprised by how much or by how little the different analyses overlapped in the choice of elements, effects and key effects?	Whether the expert expected to see commonalities.
Was there much variety in the emphasis of the analyses, and what do you think that means? For example, maybe one had more surface elements and the other more abstract effects, or maybe one chose only one or two effects as key while the other chose many.	Whether the expert thinks there is a reason for any differences between the focus of TAPT analyses (if differences exist).
Do either of the artefacts you reviewed stand out to you as especially inspired or dull? Why?	Whether either artefact stands out.
Have you any other comments about the TAPT artefacts you have seen?	Catch-all.

Table 6-5 Expert review: questions on both TAPT artefacts

The questionnaire, Participant Information Sheet and Consent Form can be seen in Appendix J: Expert Review Materials.

6.1.4 Selection and allocation of artefacts

For each interview, six artefacts were needed, two from each of the three methods. A random number generator was used to choose the six artefacts for review, and to choose the order in which the artefacts were to be presented to the expert.

Two experts were asked to review material from each task domain. Once six blind artefacts had been chosen for the first of each pair of expert reviews, two artefacts were randomly selected from that pool of six to be re-visited in the second review: one of these was always TAPT-generated, and one was always generated by one of the other methods. This was intended to ensure there was some overlap between reviews, as a way to gauge expert consensus. Due to the randomised selection of artefacts, in two cases additional pairs of artefacts were generated, such that microblog and wiki experts shared three common items, not two.

The scope of this study was not sufficient to examine every artefact produced in the comparative evaluation. Table 6-6 shows the number of artefacts available for use.

	Microblog	**Picnic**	**Wiki**
TAPT	4	6	5
Scenarios	5	5	2
UD	4	4	5
Total	14	15	12

Table 6-6 Number of design artefacts available for expert review

Note that TAPT was used to produce 6 wiki artefacts. However, the artefact produced by Group D (see Appendix G: Artefacts from the Comparative Evaluation) was very difficult to understand, and was thus removed from the pool.

The microblog artefact produced by Group I was also removed from the pool, on the basis that this design was Group I's second use of TAPT. Of course, many of the participants in the comparative evaluation who used Scenarios could have used that method any number of times in the field (in contrast to TAPT users, who all used it for the first time): one could argue that there is no need to exclude an artefact produced by participants with some prior TAPT experience. However, it is reasonable to eliminate as many variables as possible. All other TAPT-produced designs resulted from participants' first use of TAPT, so drawing from that pool of designs removed the variable of participants' level of TAPT experience.

Finally, one picnic artefact (G, produced with Scenarios) was excluded, because it answered a different task to all other picnic artefacts. The method of selecting picnic artefacts was slightly different: picnic artefacts addressed different perceived tasks, broadly falling into one of two categories: websites to support conferencing, and online shops. Artefact G, a website for organising picnics, was the exception. It was felt that differences in the core functionality of artefacts could be distracting to an expert who is trying to comment on other issues such as accessibility and creativity of artefacts, and therefore the following approach was used:

- Randomly select one picnic artefact which supports conferencing from each method
- Randomly select one picnic artefact which support shopping from each method

This resulted in six artefacts, three about conferencing and three about shopping. This approach excluded the ‘picnic organiser’ site from the pool (G, produced with Scenarios), but resulted in the ability to examine the two main types of functionality produced.

The artefacts were shown to the experts in two groups of three (that is, all the shopping artefacts together, and all the conferencing artefacts together) in order to minimise the distraction which would occur from the change of function. The order within the two groups of three was randomly selected.

The selection method for the second review was the same, except that the researcher ensured that one of each set of three artefacts was repeated from the first review, and that one of the two TAPT designs was repeated from the first review.

Nine of the 41 artefacts were generated by university participants as opposed to IBM participants. The researcher did not alter the artefact selection process to target one type of artefact over another, although the provenance of the artefacts was tracked to ensure that some coverage of both sources occurred. In the event, four experts saw two university-sourced artefacts, one expert saw one such artefact, and one expert saw only artefacts generated by IBM staff.

Appendix G: Artefacts from the Comparative Evaluation details the artefacts produced in the comparative evaluation. Table 6-7 lists the artefacts that were reviewed by each expert, in the order in which they were reviewed. The first letter shows the group which produced the artefact, and the letter in brackets show the method used (U = Unstructured Discussion, S = Scenarios, T = TAPT). For example “N (T)” on the row of Expert A (microblogs) refers to the microblogs artefact produced by Group N, who used TAPT. Artefacts in bold were produced by university participants. Full TAPT artefacts were presented in the second half of the interview in the same order in which they were shown as blind artefacts.

Expert	Artefacts
A (microblogs)	N (T), E (S), **Q (U)**, **S (S)**, F (U), B (T)
B (microblogs)	O (S), F (U), E (S), N (T), K (T), L (U)
C (picnics)	P (U), M (S), C (T), **R (S)**, **S (U)**, L (T)
D (picnics)	F (T), J (U), A (S), **S (U)**, L (T), **U (S)**
E (wikis)	K (S), **U (U)**, J (T), A (T), N (U), **T (S)**
F (wikis)	B (U), R (U), K (S), **T (S)**, A (T), G (T)

Table 6-7 Order in which artefacts were reviewed

Artefacts examined by both experts allocated to the each task domain were:

- Microblog artefacts E (Scenarios), F (Unstructured Discussion) and N (TAPT)
- Picnic artefacts L (TAPT) and S (Unstructured Discussion)
- Wiki artefacts K (Scenarios), T (Scenarios) and A (TAPT).

6.1.5 Processing the results: qualitative content analysis

Qualitative content analysis was used to process the results, guided by (Gibbs, 2004), (Strauss, 2004) and (Weiss, 1995). The following approach was taken:

1. Anonymise the audio files
2. Transcribe the interviews
3. Conduct topic coding on the data using NVivo[21], to elicit key words, concepts and categories
4. Conduct a subsequent coding to place the identified 'nodes' (coded key words, concepts and categories) in common categories: for example, in the case of blind artefacts, categorise nodes which are about innovation; challenges; replication; analysis; accessibility
5. Scrutinise the categorised results for:
    a. interesting or unexpected comments
    b. comments demonstrating strong feelings (positive or negative)
    c. agreement or disagreement between experts on common artefacts and overall themes
6. Analyse the comments of interest to elicit results
7. Examine NVivo nodes and (where relevant) numeric data for any further evidence for or against the results (for example, testing the impact of sequence on results based on comments by two experts – see Section 6.2.3)

[21] Qualitative data analysis software package available at the university.

Appendix K: Sample Expert Review Transcript) gives a sample of transcribed material while Appendix L: NVivo Screenshots) provides screenshots of some of the coded nodes.

In transcriptions, "…" is used to indicate a pause in speech and "[…]" is used to indicate omitted material.

6.1.6 Processing the results: lightweight numeric analysis

The results from this study were primarily qualitative in nature, but this did not preclude a lightweight quantitative analysis, to aid comparison of responses. Experts were asked whether each blind artefact was innovative, met the challenges they had identified in the task, replicated the experience in question, appeared to be built on the basis of an analysis, and was accessible. Answers to these specific questions were subject to a numeric analysis as follows.

Each artefact was scored based on expert answers: a score of 1 was given if the expert said 'yes, the artefact did this', 0 if they said it did not, and 0.5 if the answer was 'somewhat' or 'to an extent'. If the expert wasn't sure or felt unable to answer the question, no score was given and that particular answer was excluded from the subsequent analysis. If an expert said an artefact replicated an experience on a superficial level (but not in terms of the underlying experience), a score of 0.5 was given; if an artefact replicated an experience deeply, a score of 1 was allocated.

The scoring helped understanding and comparison of comments about artefacts with regard to their levels of innovation, meeting challenges, replication, analysis and accessibility. It also enabled comparison of whether experts:

- varied in their average rating of artefacts
- varied in their ratings over the course of examining six artefacts
- agreed on artefacts which they shared in common
- gave different scores to artefacts generated by IBM participants against those generated by university participants.

When presenting the results as numeric scores, the numbers are given to one decimal place. Although the numbers could be given to a greater level of accuracy, this lower granularity was chosen to emphasise that the results are rankings based on responses of "yes", "no" or "somewhat", itself a coarse level of granularity.

A small number of t-tests were conducted, as follow:

- Comparing scores of IBM vs university artefacts (result: not statistically significant)
- Comparing scores of TAPT, Scenarios and UD artefacts (result: UD scores statistically significantly better than TAPT and Scenarios with p=0.01 and p=0.0005 respectively; TAPT scores are not statistically significantly better than Scenarios scores)
- Comparing the analysis scores of TAPT, Scenarios and UD (no statistically significant results)

The results are discussed fully in subsequent sections.

6.1.7 Processing the results: use of mixed methods

The use of mixed research methods facilitated triangulation of results. For example, a set of qualitative comments on artefacts were coded as being ‘enthusiastic about’ or ‘disliking’ the artefacts, and the numeric scores which had been generated enabled the researcher to further explore this area, for example by searching for further comments on artefacts which had been scored particularly well or badly.

Generally, the top-rated artefacts were those about which experts were most enthused: for example, Expert A particularly praised artefacts microblog-F and microblog-Q (both from Unstructured Discussion): these were also Expert A’s highest-scoring artefacts (0.9 and 1.0 respectively, compared to an average score from A of 0.8).

However, a low rank did not necessarily mean an expert strongly disliked an artefact. For example, Expert B rated artefact microblog-E (Scenarios) the lowest of all the artefacts he saw (score: 0) and commented negatively (“The others seem to be better in my opinion.”), yet he did not appear to have a strong emotional reaction to the artefact, simply saying that it did not satisfy his requirements. By contrast, Expert A rated the artefact more highly (score: 0.7), yet like Expert B she was not greatly enthused by the artefact, although she did feel that it satisfied the aspects she was asked about. For this particular artefact, the experts’ ratings are different but their tone of comment is similar. It is possible that Expert A was simply more positive in her language about the artefact.

This apparent disparity of ratings emphasises that although the numeric ratings are a tool for approximating expert opinion, the ratings alone are a less robust indicator than when they are taken in the context of qualitative comments.

6.2 Experts and Factors to Impact Results

This section discusses factors that impact the results of the expert review. First there is a description of the experts' backgrounds and responses to the design tasks as well as their average scores of design artefacts. This is followed with a discussion of Expert C, who appeared to approach the review from a significantly different standpoint to the other experts. The section continues with a discussion on the impact of whether artefacts from the comparative evaluation were designed by IBM or university participants, and the impact on scores of the stage within each interview at which artefacts were considered. The section closes with a discussion of the extent to which experts agreed about artefacts which were shared in common, as a gauge of expert consensus.

6.2.1 Experts, their perceptions of the tasks, and their ratings

Four of the experts (Experts A, B, E and F) were academics working in the field of Computer Science (three from the University of Southampton, one from the University of Edinburgh). Expert C was a Professor of Sociology and the only expert with a non-engineering background. Expert C responded differently to other participants, discussed in Section 6.2.2. Expert D was the only expert to be currently working in industry and not academia: this did not seem to cause any difference in her approach to the task compared to the academic Computer Science experts.

The experts were shown the design task given to participants of the comparative evaluation and asked to comment on what they perceived to be the key challenges of the design task and how they would approach the task themselves. This information, along with a description of each expert, is given in Table 6-8.

Experts largely identified similar key challenges. For example, A, B, E and F all mentioned usability, while A, C and D commented on replication. E and F, who examined the wiki task, identified the same two challenges (a suitable interface and eliciting contributions).

Half of the experts described a broadly similar approach to the task: Experts A, B and F all described acquiring data from end users. Experts D and E described other approaches, while Expert C remarked that she did not know how she would approach the task.

Expert	Challenges	Approach
A (microblogs): Second year PhD student studying blogs	Usability, replicate microblog conventions	Interview, observe, consider hardware
B (microblogs): Lecturer with research interests in social networking	Understanding context, acquiring identifiers for users, accessibility	Start by gathering data on people, expected volume of transactions
C (picnics): Sociology Professor	Translating the multiplicities and dynamism of social interactions into binary code	Approach: don't know, couldn't do it
D (picnics): User experience designer with a Bachelors degree in Psychology and a Masters in HCI	Grasping what the site needs to do, and how to reconceptualise picnicking as a website. Change things to a point but be recognisable	Work out what makes a picnic experience, and which aspects you can somehow translate. Maybe bounce ideas off others
E (wikis): Final stage PhD student looking at wiki user experience and emotions in the wiki user experience	Elicit contributions, balance the interface (enable rich commentary but don't intimidate less technologically-oriented people with complexity)	Design the space, seed it with content
F (wikis): Senior Research Fellow with interests in agile methods and socially interactive designs	Elicit contributions, provide a very intuitive interface	Early design work, mock ups, get input from laypeople early on

Table 6-8 Experts and their comments on the task

Scores were calculated for each artefact which underwent review. Artefacts were scored on individual facets (level of innovation; ability to meet the design challenges identified by the expert; replication of experience; analysis of experience; accessibility). Each facet was rated at 0 if the expert felt the artefact failed to meet the facet, 1 if the expert felt the artefact succeeded in this respect and 0.5 if the expert felt the artefact 'somewhat' met the facet. (If the expert felt they did not know whether the artefact met the facet, that individual answer was excluded from the numeric ratings.) The rating of each of the five facets was averaged to give each artefact an overall score from the expert, between 0 and 1.

The scores referred to in this chapter were calculated based on the verbal answers given by experts. Table 6-8 shows the average calculated score from each expert.

As can be seen, scores vary wildly, with Expert A being most generous and Expert C the least generous. Their scorings of individual artefacts should be considered within this context.

Expert	Average score
A	0.8
B	0.3
C	0.0
D	0.6
E	0.5
F	0.2

Table 6-9 Artefact scores (by expert)

6.2.2 Expert C: an outsider?

Expert C, a Professor of Sociology, gave particularly low ratings. This is probably connected with her initial stance: she went into the interview expressing strong doubts about the possibility of ever being able to replicate a picnic experience with technology. As well as saying she did not know how she would approach the task and that she couldn't do it, she perceived the task as being about reducing social interactions to "binary code" rather than a high-level system design. At the close of the interview she talked about how she feels it is impossible to abstract a picnic because they are all so different:

> *It depends on who's there, why they're there, whether it's raining, whether it's in Mumbai or Southampton, all of these things matter in terms of what the meaning of it is… so I have a real issue with trying to abstract it.*

As well as feeling that the task was extremely difficult, Expert C appeared to view the participants of the comparative evaluation as rather different to the people she is used to interacting with, distancing herself from them. For example, she sometimes commented on the background and language of participants, for example remarking that "There's a whole other layer here about how these […] software designers […] are conceptualising their job […] their understanding of their role and relationship." When she saw the abbreviation "i8n" (internationalisation) in artefact picnic-S (Unstructured Discussion), she laughed and asked what it was, remarking "I thought it was a protocol" and asking "so you think this is like 'text language'?" By contrast, Expert D saw the same abbreviation and simply asked "Is that [pointing] 'internationalisation'?"

In summary, Expert C was sceptical about whether the task was possible, and appeared to view herself as something of a 'foreigner' in the context of designs built by engineers. Her low ratings of artefacts perhaps reflect this scepticism.

Expert C's responses are included in the results presented in this chapter.

6.2.3 The impact of artefacts' provenance and sequence on scores

It is possible that artefacts generated by IBM and university participants might show different qualities, and also that artefacts rated early or late on in the interview process might be rated differently. This section discusses these two areas.

Table 6-10 shows the relative scores given to artefacts generated by IBM and university participants. Some variety is evident, but the difference is not statistically significant.

	IBM	**University**
Mean	0.39	0.47
Standard deviation	0.31	0.45
Variance	0.09	0.21

Table 6-10 Statistics on scores, by participant background

University artefacts > IBM artefacts (t=0.47. df=34, not significant)

Experts were noticeably enthused by some artefacts and disenchanted by others, shown in Table 6-11. Artefacts from university participants are shown in bold.

	Enthusiasm	**Disenchantment**
Microblog	F, L, **Q, S**	
Picnic	**U, S**	**R**
Wiki	A, N, **U**	J, K, **T**

Table 6-11 Artefacts to evoke strong feeling from experts (comparing IBM and university)

Just over a quarter (8 out of 29) of the artefacts used in the expert review were designed by university participants. As can be seen, university participants generated five of the eight artefacts to prompt enthusiasm (although only three individual groups produced the five artefacts) and two of the four artefacts which were strongly disliked. Although these numbers aren't what might be expected, in the context of the small numbers being considered here the difference is not striking, especially given that only three separate university groups produced the well-received artefacts.

Two experts commented on differing perceptions over time, with Expert C saying "You'll get different responses to that question as you ask it six times" and Expert F remarking later on in the interview that she thought she was harder on artefacts at the start. Table 6-12 shows the average score given to each blind artefact in chronological order.

First	Second	Third	Fourth	Fifth	Sixth
0.2	0.6	0.4	0.5	0.4	0.5

Table 6-12 Mean artefact scores from all experts (by time)

As can be seen, there is no particular trend for ratings to improve or worsen over the course of time. Although Expert F felt she was harder on artefacts early on, her specific ratings do not suggest this, consisting of the sequence: 0.1, 0.3, 0.5, 0.0, 0.0, 0.4.

6.2.4 Expert agreement

Each pair of experts examined two or three artefacts which were the same. This was primarily to establish whether their expert view was substantiated: many disagreements between experts could imply that their knowledge was variable.

Experts disagreed on a total of three artefacts, as follows:

Microblog experts shared three common artefacts (Table 6-13), and did not strongly disagree about any of them. Expert A was more generous than Expert B in general (average rating of 0.8 compared to 0.3), which accounts for the disparity of their ratings of artefacts microblog-F and microblog-N. The numbers suggest the two experts strongly disagree about the quality of artefact microblog-E, but their comments on the artefact are largely in agreement, as discussed in Section 6.1.7.

Artefact	Expert A	Expert B
microblog-E (Scenarios)	0.7	0.0
microblog-F (UD)	0.9	0.5
microblog-N (TAPT)	0.8	0.4

Table 6-13 Ratings of repeated microblog artefacts (by expert)

The picnic experts examined two artefacts in common, picnic-L (TAPT) and picnic-S (UD). The picnic experts had similar responses to artefact picnic-L, but disagreed on artefact picnic-S, probably because they interpreted the task differently (discussed in Section 6.5) and artefact picnic-S answered the task as interpreted by Expert D but not the task as interpreted by Expert C.

Artefact	Expert C	Expert D
picnic-L (TAPT)	0.1	0.3
picnic-S (UD)	0.0	0.9

Table 6-14 Ratings of repeated picnic artefacts (by expert)

The wiki experts looked at three common artefacts, wiki-A (TAPT), wiki-K (Scenarios) and wiki-T (Scenarios), and disagreed on two of these, artefacts wiki-A (TAPT) and wiki-K (Scenarios).

Artefact	Expert E	Expert F
wiki-A (TAPT)	0.9	0.0
wiki-K (Scenarios)	0.1	0.5
wiki-T (Scenarios)	0.0	0.0

Table 6-15 Ratings of repeated wiki artefacts (by expert)

The experts strongly disagreed on artefact wiki-A (TAPT): Expert E very much liked it, describing it as his "favourite" artefact and "a brilliant idea" which "embraces the wiki concept" and is "good for a museum exhibit". By contrast, Expert F said it was very light on detail and had "no big ideas": she felt it "could be talking about anything" and that "there's not actually that much there that displays an understanding of the context of the domain."

From these comments, it would seem that Expert E felt the idea behind the artefact was very good indeed, but that Expert F found the artefact hard to understand due to its brevity, leading her to conclude the designers did not understand or respond to the task.

The experts also disagreed over artefact wiki-K (Scenarios): Expert F gave it an average score of 0.5, while Expert E rated it at an average of 0.1. Their comments are not dissimilar, however: Expert E talked about the group's focus on viewing not editing and assuming visitors were comfortable with technology. Expert F described it as unexciting and said it relied on "gadgets" for accessibility. The two only strongly disagreed on two aspects:

1. Replication of wikis. Expert E said "I don't think it really gets the crux of what a wiki is needed there for", while Expert F said it was "very wiki-like" because its random selection of data was analogous to a wiki's "thing of the day".
2. Meeting challenges (usability; eliciting contributions). Expert E felt that "in terms of an editing interface it's not very good" and that it "makes assumptions about people's comfort with technology". By contrast, Expert F said "they've thought about the kind of technology that'd help them with this – which is the touch-screen or wii type access."

From this it would appear that the experts disagreed on whether the artefact truly re-provided wiki experiences or used suitable technology for the task.

In summary, experts disagreed about three of the eight artefacts which were cross-examined:

1. The picnic experts disagreed on artefact picnic-S (Unstructured Discussion), probably because it included shopping elements which Expert C felt was not part of the task and which Expert D felt were appropriate.
2. The wiki experts disagreed on artefact wiki-A (TAPT): Expert E liked the idea, Expert F found the design difficult to understand.
3. The wiki experts disagreed on artefact wiki-K (Scenarios). Expert E felt it did not understand wikis, Expert F felt that it did.

The first disagreement seems to stem from different interpretations of the task, while the second is apparently rooted in a lack of clarity within the artefact. The final item appears to be a fundamental disagreement.

The fact that the pairs of experts largely agreed on artefacts suggests that their opinions are substantiated and can be taken seriously.

6.3 Results: Blind Artefacts

This section presents results related to the blind artefacts presented to the experts. Section 6.3.1 describes the results and Section 6.3.2 discusses the impact normalisation may have had upon responses to artefacts.

6.3.1 Scores and comments

As can be seen in Table 6-16, experts rated Unstructured Discussion artefacts more highly than those produced by TAPT or Scenarios (a more detailed breakdown is given in Appendix M: Numeric Expert Review Data). TAPT's lead over Scenarios was not statistically significant, although the lead of Unstructured Discussion-generated artefacts over artefacts from both other methods was shown to be statistically significant.

	TAPT	**Scenarios**	**UD**
Mean	0.38	0.29	0.58
Standard deviation	0.34	0.38	0.45
Variance	0.12	0.14	0.20

Table 6-16 Statistics on ratings

UD > TAPT ($t=2.61$, $df=110$, $p=0.01$); UD > Scenarios ($t=3.67$, $df=115$, $p=0.0005$); TAPT > Scenarios ($t=1.23$, $df=108$, not significant)

Appendix M: Numeric Expert Review Data lists ratings in response to specific questions. Most responses follow the broader trend of Unstructured Discussion being rated highest, followed by TAPT then Scenarios. The only exception was that Scenarios artefacts were rated as slightly more innovative than TAPT artefacts and that Scenarios and TAPT artefacts were rated equally on use of analysis.

Regarding analysis, the average score of Unstructured Discussion artefacts was 0.5, compared to 0.3 and 0.2 for TAPT and Scenarios artefacts respectively. Unstructured Discussion's rating is not statistically significantly greater than that of TAPT or Scenarios (Table 6-17), but it is nonetheless striking that the two methods which enforced a level of analysis were both rated worse than Unstructured Discussion, which prompted no analysis. This probably relates to the issue of brevity of artefacts, which is discussed in Section 6.3.2.

	TAPT	**Scenarios**	**UD**
Mean	0.25	0.21	0.50
Standard deviation	0.35	0.33	0.48
Variance	0.13	0.11	0.23

Table 6-17 Statistics on ratings of analysis

UD > TAPT (t=1.42, df=20); UD > Scenarios (t=1.67, df=22);

TAPT > Scenarios (t=0.27, df=20); no significant results

Numeric ratings are not the only way to gain insight into the artefacts and do not necessarily indicate strength of feeling. A complementary approach is to examine which artefacts prompted strong language (whether positive or negative) as an indicator of the quality of an artefact. Table 6-18 lists the artefacts to provoke such commentary.

As can be seen, one TAPT and two Scenarios artefacts provoked strong positive comments, in contrast to six Unstructured Discussion artefacts. One TAPT and three Scenarios artefacts provoked strong negative responses, while no Unstructured Discussion artefacts did this. Clearly, Unstructured Discussion artefacts were well received. Two aspects may have affected the ratings:

Firstly, participants of the comparative evaluation who used TAPT appear to have struggled to complete the task in the given time. Responses in Chapter 5 discussed this. Participants had an equal amount of time to use all three methods, but for Scenarios and particularly for TAPT (which was entirely new to participants) some of that time had to be spent on understanding the method and applying it. This is discussed in Section 6.4.1.

Secondly, the normalisation process by which the blind artefacts were produced may have left TAPT and Scenarios artefacts at a disadvantage. This is discussed in the next section.

	Enthusiasm	**Disenchantment**
TAPT	1. Expert E said wiki-A was his "favourite" (described in Section 6.2.4)	1. Expert E was "angry" about wiki-J allowing visitors to record over one another's comments
Scenarios	**1.** Expert A liked artefact microblog-S, repeatedly referring to it as "cool" **2.** Expert D was enthused by artefact picnic-U, finding it "innovative", "fun" and able to balance business and experiential aspects	1. Expert C strongly disliked artefact picnic-R calling it "shocking" and "a pretty poor, thoughtless account" 2. Expert E found artefact wiki-T "annoying", saying "it isn't a wiki" and "it seems confused" 3. Expert E found artefact wiki-K "especially dull"
Unstructured Discussion	1. Expert A and B liked artefact microblog-F: Expert A repeatedly called it "cool" and Expert B said "I like this one" and "it's not electronic, but it does address what I consider to be important requirements" 2. Expert A was also enthused by artefact microblog-Q: "I would definitely do something like this" 3. Expert B described artefact microblog-L as his "favourite" artefact 4. Expert D "liked the idea" of picnic-S, which was immersive and balanced business and experiential aspects 5. Expert E found artefact wiki-N "novel" and "intriguing", saying it had "grabbed what a wiki's about" 6. Expert E liked wiki-U, saying it was "really insightful" and that "they really know what they're talking about in terms of wikis and web 2.0"	

Table 6-18 Artefacts which provoked strong comments

6.3.2 The impact of normalisation

Expert B described three artefacts as structured: artefacts microblog-E (Scenarios), microblog-F (Unstructured Discussion) and microblog-K (TAPT). This spread across all methods suggests the normalisation of artefacts was successful in masking their provenance.

However, experts sometimes commented on finding blind artefacts hard to understand, as follows:

- Expert A, artefact microblog-B (TAPT): she read parts aloud, asking "if there are public screens how do you consume stuff? If it's voice snippets or templates, templates is ok but how do you consume voice snippets, all at the same time? […] So they listen… public screens say someone's said something, then you can use your handheld device to listen?"
- Expert B, artefact microblog-K (TAPT): "It's inconcrete. […] There are a lot of things that are not clear. "
- Expert B, artefact microblog-E (Scenarios): "how do you get messages that are relevant, how do you know if someone left you a message […] how do you be alerted to someone going and doing some activity unless you switch on a device…."

None of the above artefacts were generated with Unstructured Discussion, strongly suggesting that artefacts generated with UD were clearer to experts.

One can also consider ratings of methods: sometimes an expert would answer a question with "I don't know" or "I can't tell", because the artefact was unclear. If all questions were answered, the expert review would result in 60 observations from artefacts generated with each method (5 observations per artefact, 12 artefacts assessed per method). In the event, there were 59 observations from artefacts generated with Unstructured Discussion, 58 from those generated with Scenarios and 53 from those generated with TAPT: TAPT artefacts appeared less clear.

Table 6-19 shows the average word count of the normalised artefacts, by method. It can be seen that TAPT artefacts were much shorter than the other two, and that Unstructured Discussion artefacts were the longest.

Method	**Average word count**
TAPT	69
Scenarios	105
UD	129

Table 6-19 Word counts of normalised artefacts (by method)

Other comments concerned a lack of depth or thoroughness of blind artefacts. Expert C found all artefacts superficial, and the other experts commented as follows:

- Expert D found the following artefacts “basic”, “short” and with “not much detail”: picnic-A (Scenarios), picnic-F (TAPT), picnic-L (TAPT).
- Expert E, artefact wiki-K (Scenarios): “Method of input is not clear.”
- Expert E, artefact wiki-A (TAPT) “it doesn’t come up with anything about how you’d input graffiti and the interface.”
- Expert E, artefact wiki-J (TAPT): “it’s ambiguous over what they’ll use in terms of input.”
- Expert E, artefact wiki-U (Unstructured Discussion) felt they mentioned accessibility but did not address the issue: “They make a specific point about accessibility […] They don’t really mention anything about visual accessibility, text size etc.”
- Expert F, artefact wiki-R (Unstructured Discussion): “technology-free.”
- Expert F, artefact wiki-A (TAPT): “very light on any kind of detail.”

Some of the above artefacts are from Unstructured Discussion, but most (five out of the nine in the above list) are from TAPT. Like the perceived lack of clarity, the lack of depth is probably linked with the brevity of the TAPT artefacts.

In two instances signs of the method used crept through and led to positive comments:

1. Expert A liked the “story-style” presentation of artefact microblog-S (Scenarios).
2. Expert F liked that in artefact wiki-G (TAPT) they “set out in their design criteria that they wanted something visually engaging, simple and intuitive to use, approachable.”

The microblog artefact produced by group S did not include a design description, only one scenario and one persona. Consequently, in the normalisation process the author added some context to the scenario from the persona (changing “Frank” to “Frank, a care home resident”) and presented the scenario as the blind artefact. Expert A clearly found the presentation style engaging.

Expert F comments positively about the ‘design criteria’ in artefact wiki-G: these ‘criteria’ were where the group wrote what they perceived to be new experienced effects of their design. This evidence of analysis was greeted very positively as a sign of understanding (“there was some analysis there […] so they’ve shown [understanding]”).

It would appear that any markers that analysis was conducted evoked positive comment: For example, Expert B’s first comment when reading picnic-L (Unstructured Discussion), later identified as his favourite artefact, was “Oh!

‘Assumptions’”, quoting their initial heading. He felt that analysis had been conducted.

In contrast, as described above, many participants disliked the apparent lack of analysis in the shorter artefacts. For example, Expert D said of artefact picnic-L (TAPT): “It’s incredibly short, which doesn’t suggest to me that huge amounts of thought have gone into it.”

The process of normalising the artefacts involved removing the aspects of TAPT and Scenarios which were characteristic of those methods: the Teasing Apart analysis of TAPT, and the personas and scenarios generated with Scenarios. Excluding this information from the TAPT and Scenarios artefacts appears to have led experts to a) struggle to understand them and b) feel the artefacts lacked depth. For example, upon seeing all of artefact wiki-A (TAPT), Expert F said she felt the blind artefact was light on detail because that detail was implicit in its analysis. It seems likely that reduced Scenarios artefacts also suffered in this respect, although as experts were not later shown full Scenarios artefacts this can only be speculated on.

In summary, Unstructured Discussion produced highly-rated artefacts. This would appear to be because the blind UD artefacts were clearer to experts, and because TAPT and Scenarios artefacts did not convey the full richness of their analysis. In effect, blind UD artefacts were better presented than blind TAPT or Scenarios artefacts, which interfered with expert analysis of the actual ideas. The reduction of artefacts for purposes of normalisation appears to have done them a disservice.

It is also possible that UD artefacts were more highly rated as they recreated experiences in a more obvious, superficial way; while TAPT and Scenarios may recreate experience in a more sophisticated way. This would mean that when an expert examines a UD design they may more easily see how that design captures obvious elements of the original experience. An expert might not fully understand the way in which a TAPT artefact replicates an experience until they see the full TAPT analysis. For example, a 3D virtual cracker more obviously captures the experience of pulling a Christmas cracker than a more nuanced experience as described by Dix (see Section 4.1), even though the latter may actually deliver an overall experience which is quite similar to pulling traditional crackers.

After reviewing six blind artefacts, each expert was shown the full body of the two TAPT artefacts which were included amongst the blind designs. The results from this aspect of the expert review are given in the next section, which sheds further light on expert perceptions of the artefacts, particularly with respect to the relationship between the initial analyses produced in the comparative evaluation and the resultant designs.

6.4 Results: TAPT Artefacts

This section discusses responses to the full TAPT artefacts, the Teasing Apart analyses and Pieced Together designs. After a discussion of the brevity of the artefacts, this section describes comments on the strength of TAPT analyses, discusses an apparent disjunction between Teasing Apart and Piecing Together, and finally touches upon the possible leading nature of example artefacts.

6.4.1 Brevity of TAPT artefacts

Three experts commented on the brevity of TAPT analyses and designs:

- Expert B reported thinking that artefact microblog-K must have more material than he saw, as the design was so brief. He remarked of the analysis, “It’s as if these people had less time.”
- Expert C remarked of artefact picnic-C “It’s like they got bored or in a hurry and had to finish fast,” and of artefact picnic-L that they stopped after identifying items rather than developing their ideas.
- Expert D speculated that design artefact picnic-F was “one of the simplest ones” because the Teased Apart analysis was “implicitly already on the sheet” which participants of the comparative evaluation handed in. She felt that they described experiences in the analysis and therefore felt they did not need to repeat that in the design. She also remarked of the Pieced Together design: “It seems like a summary, not complete justification – they didn’t have a lot of room.”

These comments concern a perception that designers using TAPT did not have sufficient time or (in the case of Expert D’s comment) room to write. TAPT artefacts were especially short in comparison to artefacts from the other two methods: in addition to the issue of truncation discussed in Section 6.3.2, this is probably connected with participants of the comparative evaluation requiring more time to learn and apply TAPT (see Section 5.4.7). The above comments seem to substantiate this.

6.4.2 *Strong analyses*

A recurrent theme through comments about the TAPT artefacts was positivity about the strength of the analyses for understanding people's motivations and feelings:

- Expert A said of artefact microblog-N: "They've actually considered all these things […] it definitely made sense and they definitely put considerations into […] what happens, and why people use it."
- Expert A said of artefact microblog-B: "[The analysis] basically says a lot about how people use it and what they feel about it."
- Expert C said of artefact picnic-L: "I'd have liked them to articulate this [analysis] more […] [the abstract effects] are good in my view."
- Expert D said of artefact wiki-F: "They considered things like the emotional aspects, the abstract aspects."
- Expert E said of artefact wiki-J: "The Teasing Apart is actually very good."
- Expert E felt that that he'd probably produce a "similar" analysis to that of artefact wiki-A: "if [his] creative skills allowed."
- Expert F said of artefact wiki-A: "Well, we have a reasonable analysis here."
- Expert F said of artefact wiki-G: "They've certainly analysed what a wiki is."

No expert felt any Teasing Apart analysis had failed, although Expert C said of picnics-C that the analysis "isn't great". This result links with statements from participants of the comparative evaluation, who felt the analytical aspects of TAPT were a strength of the method (Section 5.4).

6.4.3 *Disconnect between Teasing Apart and Piecing Together*

Many experts commented on a disconnect between the Teased Apart analyses and the Pieced Together designs. The six experts each reviewed two TAPT artefacts, and eleven of the twelve reviews commented that in some way the design did not fully reflect the analysis (quotations are in Appendix N: Expert Review Quotations). Many, although not all, of the comments focused on analyses mentioning emotional aspects and designs failing to address these aspects, and analyses focusing on feelings while designs focused on functionality. This links with the observed discomfort designers had in analysing emotions and abstract meanings, commented upon by three experts:

- Expert B noted an emphasis on surface and literal elements in both microblog TAPT artefacts, suggesting maybe this was as the work was done by people with a background in Computer Science ("If you do the trial in [the School of Electronics and Computer Science], [we] will tend to place more emphasis on this, we don't spend time to make the case for what is the experience"). He similarly commented about how although artefact microblog-K's designers have user-oriented categories, they talk about technology, touching on emotions but not exploring them.
- Expert C notes that artefact picnic-C Teased Apart superficial aspects (such as food) but found it hard to unpack meanings or shared social experiences, saying "They find it hard to grasp/talk about the social" and "They cop out."
- Expert F felt that both groups found it hard to think about experienced effects: "Probably because they're engineers, they're grappling with the experienced effects and what they may mean. That seems to be quite alien with them. So I think they're kind of coming up with things off the top of their head. They don't necessarily know how to think about it."

Similarly, some experts noted an emphasis on surface elements and literal effects:

- Expert B observed this about both of his TAPT artefacts, speculating it was because the designers were Software Engineers who are more comfortable with the tangible.
- Expert D observed this focus in both her TAPT artefacts.
- Expert F also felt that designers of both TAPT artefacts had found it difficult to think about both abstract effects. Like Participant B, she speculated that this was "As they're engineers, they don't necessarily know how to think about it."

Two comments concerned design artefacts including facets not evident in the initial analyses:

- Expert C noted that in artefact picnic-C the analysis did not emphasise anticipation, yet the design was "dominated" by it.
- Expert F noted that "the only really imaginative thing" in artefact microblog-A was "this idea of the wall growing over time", yet there was nothing about this in the analysis.

In general, it appears there was a strong feeling that analyses were stronger than designs, and that designs were not entirely informed by analyses – particularly regarding abstract and emotional aspects. An improvement to the method would be to encourage users to mark in their designs where they feel the key experienced effects manifest themselves, prompting them to be more explicit about how they feel they are replicating experiences.

6.4.4 Leading examples

The preparatory work conducted before the comparative evaluation (Section 5.1) included the task of choosing appropriate examples of TAPT's use. The intent was to adequately illustrate the nature of each stage of the TAPT process without suggesting possible solutions to the allocated design tasks.

Participants in the comparative evaluation did not comment on the example artefacts. However, some TAPT analyses do appear to have been 'led' by those examples, in that they include material from the sample analyses. Two experts remarked upon feeling surprised by 'odd' experienced effects:

- Expert B, artefact microblog-K, uncertainty about audience and responses: "I'm surprised they chose this because I'm not clear about it."
- Expert B, artefact microblog-N, anticipation: "I'm not sure what it means."
- Expert D, artefact picnic-F, cultural connotations: having been asked if she found any effects surprising or odd, she replied "'Cultural connotations' was an interesting one. So yes, I guess [I did]!"

The above three effects were present in the example TAPT artefacts given to participants of the comparative evaluation. The lack of comments upon feeling led by the examples suggests that participants were unaware of the phenomenon, but their usage of these pre-provided effects in a way that puzzled the experts (i.e. was inappropriate), suggests that some amount of leading did occur.

6.5 Results: Broader Insights

This section discusses insights gained from expert comments at a broader level than that of individual artefacts. Section 6.5.1 describes comments on the nature of the picnic and wiki tasks and Section 6.5.2 discusses a comment on the way TAPT is applied.

6.5.1 Insights into patterns of picnic and wiki designs

Expert comments shed light on aspects of responses to the picnicking and wiki tasks.

Expert C was disappointed to see shopping and marketing aspects in the picnic artefacts. For example, of artefact picnic-R (Scenarios), Expert C said "Well, it's a shopping website, isn't it?" and later reflected "[It] was pretty shocking […] that was a pretty poor, thoughtless account." Expert D was not shown that artefact, but in general she commented positively on artefacts that did involve business aspects, for example of artefact picnic-A (Scenarios): "They've thought about the business side, whereas the other two hadn't. That's clearly something the company's interested in, that pays you to build the website, so that's a good thing." Later, she added: "I think probably food has to play a part in it because they're trying to sell picnic products, so there's a certain amount of meeting the marketing needs." At this point, the researcher asked, "Do you see the task as partly to do with selling picnic stuff?" to which she said, "Yes. Or at least using the picnic stuff you'd be able to sell to them."

This represents a very strong difference in how the experts fundamentally framed the task. This difference was also evident in the comparative evaluation. In the evaluation, fifteen picnic artefacts were generated: nine online conferencing websites, five shopping websites, and one organisational tool. The author's initial response to this variety was that participants had failed to understand the task or had 'given up' (Section 5.3.2), but the contrasting attitude of the two experts implies that there is more to it than this. It would appear that the picnic task was not as clearly framed as the other two tasks, and perhaps that participants were strongly affected by their professional context and driven by commercial experience.

Expert comments also clarified an aspect of responses to the wiki task. The plethora of wall-based interfaces was puzzling (Section 5.3.3). The two wiki experts also noticed this: Expert E described a theme about "the wall as an interactive space", and felt the wall-related theme was a good reflection of the Berlin Wall, which was interactive in terms of graffiti: he liked the use of graffiti walls in artefacts wiki-A (TAPT) and wiki-N (Unstructured Discussion). Similarly, Expert F expressed that it was good that most design teams realised that "You'd actually want to have in some kind of way to have a big screen."

This suggests the experts expected (and approved of) wall-based interfaces, and also that the end domain of a museum about the Berlin Wall led participants towards wall-based interfaces.

6.5.2 Miscellaneous comments

Finally, Expert F approached the author the day after participating in the expert review with further comments. She reflected that in a design exercise there is a need to understand the domain of implementation as well as the technology being used, and suggested that asking participants to analyse technological facets but not the domain might be a weakness. She referred to Participatory Design (Schuler, 1993), which includes experts from both fields, as a precedent in the area, and suggested that following that precedent might be a way to obtain richer, more innovative results.

6.6 Conclusions

This chapter presented the methodology and results of an expert review of the design artefacts produced in a comparative evaluation of TAPT, Scenarios and Unstructured Discussion. Six experts were recruited, one of whom was not from an engineering background. Her responses were markedly different from those of the others, indicating that future reviews should take greater care over recruitment. It would seem wise to either recruit entirely from experts in a single domain of work, or if consulting a broader set of experts, including more than one from each background to gain insight into their responses and consensus.

Another lesson from Expert C's approach was that her lack of familiarity with the end domain of the picnicking artefacts (web-based technologies) left her uncertain when answering some questions. It appears that it is necessary to ensure that experts participating in such reviews are familiar with the end domain of reconstructed experiences as well as the start domain.

Results from the expert review suggest there was no great difference in the quality of artefacts submitted from IBM participants, compared to university participants. Artefacts produced with Unstructured Discussion were perceived as best by the experts, but there is strong evidence that this was because of the normalisation process, which lost some detail from TAPT and Scenarios artefacts. This means that UD was the best method for presenting ideas, but not necessarily the method to produce the best ideas. Additionally, it may be that UD artefacts recreated experience in a more obvious, superficial manner, which may have also impacted expert assessment.

It would appear that (at least when using TAPT for the first time) designers need a greater amount of time than when using other methods. One improvement to future trials might be to have participants apply TAPT once in a ‘trial run’ to begin to understand the method.

Experts perceived a disconnect between Teased Apart analyses (seen as strong) and Pieced Together designs (seen as weak), particularly with respect to abstract and emotional aspects. This indicates that TAPT needs further work on the Piecing Together phase to ensure that analyses are better translated into new designs. One possible approach might be to encourage TAPT users to mark in their Pieced Together designs exactly where they feel they are supporting the replication of abstract effects identified in their Teased Apart analyses, and to explain which ones they intentionally left out.

Expert responses shed light on two issues which arose in Chapter 5, suggesting that the two broad ‘types’ of picnic artefact were not due to some participants ‘giving up’ so much as interpreting the task differently, and that wall-based wiki artefacts were always to be expected (and were a positive thing) in the context of a wiki task oriented around the Berlin Wall.

Responses also suggested that at least two TAPT analyses were somewhat ‘led’ by the example artefacts given when first learning to apply TAPT. Much time was spent attempting to provide helpful but non-leading examples (Section 5.1, Section 5.2.3), and generally this seems to have worked (the other six TAPT analyses did not provoke comment). Including a ‘trial run’ with TAPT in future trials would not only familiarise participants with the method but might help them to begin to distance themselves from initial examples.

A comment from Expert F was of particular interest: this concerned the possibility of encouraging participants to analyse the domain of implementation as well as the technology being used, in a reflection of approaches in Participatory Design.

One of the motivations for this work was to assess whether users of Unstructured Discussion were over-confident in assessing their artefacts as being of a high quality. It would appear that the experts, like participants of the comparative evaluation, tended to rate Unstructured Discussion artefacts highly. As discussed, two factors may have impacted this: insufficient time for using TAPT, and disadvantaging TAPT and Scenarios artefacts with the normalisation process.

The following changes could be made to experiments to assess the efficacy of design methods:

- When testing new processes such as TAPT against established methods, allow participants to conduct a trial run of the new method first, to familiarise themselves with it.
- Apply a less naïve normalisation process to design artefacts when making them blind. Simple headings like 'analysis' followed by plain text of scenarios or experiential analysis would only somewhat mask the method by which an artefact was built, but would be a finer-grained tool than simply excluding some of the artefacts' material.
- Taking care when selecting experts: either ensure that the individual's expertise encompasses the technology and its application domain, or that the group as a whole is able to offer this level of insight. When seeking to verify opinions by recruiting multiple experts, ensure that the expertise of paired experts matches.

The following revisions might improve results from TAPT:

- Take steps to reduce the disconnect between Teasing Apart and Piecing Together, for example by encouraging designers to indicate in designs where they have included key effects from analyses.
- Encourage design teams to analyse the end domain as well as the start domain.

The first of the above two changes, explaining the experience-oriented reasoning behind designs, is an approach which might help overcome technology-orientation to the exclusion of users or experiences in software design processes in general.

Chapter 7 Case Studies

Chapter 4, Chapter 5 and Chapter 6 present examples and analyses of TAPT's use in controlled experiments. The type of experiment described in these chapters offers the advantage of being able to control variables (such as the tasks to which TAPT is applied) and collect a broad set of results. The disadvantage is that the experiments are artificial in nature: participants apply TAPT in laboratory-like conditions to pre-determined tasks.

In order to consider TAPT's use in a more realistic setting, the author conducted four case studies of TAPT's use in the field by professionals, presented in this chapter. This alternative approach to understanding TAPT meant the environment was not under the control of the researcher and the participants were free to use the method in the way which seemed most suitable to them. These case studies were exploratory in nature, but the objective was to understand *how* TAPT would be applied, and to find any properties of TAPT that participants felt were particularly helpful or unhelpful.

Two case studies involve TAPT as a design tool: a small software development team at IBM used TAPT to design a location-based social awareness system in a corporate environment (Section 7.2) while another IBM Software Engineer used it to re-design social networking websites into the domain of voice websites (Section 7.3). The first of these studies also involved TAPT as an evaluation tool.

The other two case studies show TAPT as an analytical tool in a research environment. The first saw two researchers at the University of Southampton use TAPT to analyse and understand genres of game (Section 7.4), while the second involved a researcher at the University of Bergen applying TAPT towards analysing and understanding location-based social services (Section 7.5).

This chapter opens with a discussion of the experimental design (Section 7.1), followed by a description of each of the four case studies. The individual case study descriptions are followed by an analysis of all of the results: this analysis is presented as a whole (rather than in four parts) in order to correlate multiple comments across the case studies. The analysis begins in Section 7.6, which discusses the motivations and expectations of case study participants and how TAPT met their expectations. Next, Section 7.7 discusses how TAPT was used in the different case studies, before Section 7.8 provides participant comments on their experiences with TAPT.

7.1 Experimental Design

The purpose of the case studies was to see how TAPT would be used by professionals in their own workplaces, in the context of their own tasks. As such, a very hands-off approach was taken by the researcher.

Participants were equipped with information about how to use TAPT and, if they wanted it, assistance in setting up studies. To avoid unnecessarily influencing proceedings, the researcher did not get involved in decisions about how to use TAPT, and merely provided information about how it could be used. Appendix O: Case Study Materials) gives the Participant Information Sheet and Consent Form for this set of case studies.

Verbal interviews were conducted with participants at the beginning and end of their work, and involved open-ended questions about (beforehand) their plans and expectations and (afterwards) the results of their work and their perceptions of TAPT's usefulness. Full copies of questions are in Appendix O: Case Study Materials. The questions did not delve into finer facets of understanding or replication, as this had been addressed in prior studies (Chapter 5 and Chapter 6). Interviews lasted ten to thirty minutes. A sample transcript is in Appendix P: Sample Case Study Transcript.

As with the expert review (Chapter 6), a semi-structured interview format was chosen as this enabled the researcher to acquire comments on consistent topics across multiple case studies (as participants answered the same broad questions), but also to follow up on interesting comments that were not anticipated in advance. Again, questionnaires were dismissed as they did not afford the flexibility to delve deep, and focus groups were not used because it was likely that more independent data would be afforded by talking to participants individually.

Participants in three of the studies chose to run studies of their own. For clarity, this chapter refers to participants in these studies run within case studies as ‘subjects’. The participants in the case studies who used TAPT and were interviewed about this are referred to as ‘participants’.

The analysis process was as follows: the anonymised audio interview data was transcribed, and answers were grouped by the question they were in response to. These were then analysed for patterns. Responses were divided into various types: responses on how the case study proceeded, how TAPT was used, and ways in which TAPT was or was not useful.

Note that in two of the studies (Spoken Web and Location-based Tools) the researcher was involved in the analysis work. (In the first study, helping Participant F Piece Together the TAPT analyses generated by the focus group, and in the second study, helping Participant I conduct an analysis of the TAPT output generated by the two focus groups.) This involvement was helpful in that the author was able to bring a strong knowledge of TAPT and its use to the table, but is likely to have altered the outcomes of each individual piece of work, and introduced a level of bias into those results. The decision was made to become involved for ethical reasons, as to do otherwise might jeopardise participants’ successful completion of their work.

7.2 Study 1: Blapr, Location-based Social Updates in the Office

This case study involved a location-based social service for the corporate world. TAPT was used to facilitate the design of the system, and as part of the evaluation process.

The ‘Blapr’ system was built by a team of five graduate recruits at IBM Hursley, who were given Friday afternoons for nine months to work on the task, from March to November 2010. The researcher acted as a ‘client’ to the team, which involved specifying the initial task and goals. The task was to provide corporate social networking via varied channels such as the web, instant messaging clients and Twitter, enabling information access in an inclusive manner. The aim was to showcase technology and skills at IBM Hursley.

The team was given access to the researcher's previous work on multimodal messaging infrastructures, particularly the example social infrastructure presented in Chapter 2. This material was made available for them to ensure that they were aware of prior work in the area. The code was available for their use if they wished, but they chose not to use it. This decision was based on a feeling that the material was relevant but more far-reaching than the specific nature of their work.

The primary goals of the project were for the team to produce a working social tool and trial its use within the IBM laboratory for at least a week with at least five volunteers. The team was asked to use Scenarios with TAPT in the design phase, because results from the Comparative Evaluation (Chapter 5) suggested that the two methods may be complementary.

The team was also asked to use TAPT as a component of their final evaluation. By Teasing Apart the experience of using their system and comparing that analysis to their initial analyses, they would be able to see to what extent they had replicated the original experiences.

7.2.1 Approach taken by the team

The team started by brainstorming different modalities by which information could be passed, before being given an alphabetically-ordered list of available technologies for use, such as ambient indicators, calendars, door displays and internal IBM systems such as BluePages (an employee directory) and the Hursley map system. (This list was alphabetically-ordered to avoid any implication of preference from the researcher.) The team then applied Scenarios and TAPT, before designing, building and testing their system. The team was given descriptions of Scenarios and TAPT that matched those given to participants of the Comparative Evaluation (Chapter 5).

The team chose to use a whiteboard when applying Scenarios and TAPT. They reported having no trouble applying either method, although they noted some difficulty in working out what granularity of persona to create (for example, the functional role of 'IBM employee' compared to a character with a name, job description and personal life).

The Blapr team considered six personas (employee, employee with special access privileges, customer, and any of the above with disabilities) and three scenarios (find a meeting, find a person, find a route to a location). They applied TAPT over the course of two Friday afternoon sessions, in the first session Teasing Apart Google maps, and in the second session Teasing Apart and Piecing Together a signpost and a tour guide. They worked on a whiteboard, with one person writing and everyone contributing words towards the lists of elements and effects in Teasing Apart, before moving into a discussion phase for Piecing Together, where they worked to find the right wording.

The team then moved on to an implementation phase, informed by the artefacts generated in the Piecing Together phase.

Unfortunately, the Blapr team was unable to conduct the proposed user-led evaluation of the system at the end of the project, as they ran out of time. They did conduct a TAPT-based evaluation (Section 7.8.4).

Interviews were conducted with the five members of the team in April 2010, shortly after the initial application of TAPT, and at the close of the project in November 2010.

7.2.2 TAPT output

The team Teased Apart Google maps, and applied both Teasing Apart and Piecing Together to the use of tour guides and signposts: they chose these experiences as they felt they were relevant to their task, which involved providing maps and routing information. At the end of the case study, one member of the Blapr team Teased Apart the Blapr solution for evaluative purposes. Figure 7-1 and Figure 7-2 show the team's application of TAPT to tour guides. Copies of material produced by the team are in Appendix Q: Case Study Artefacts.

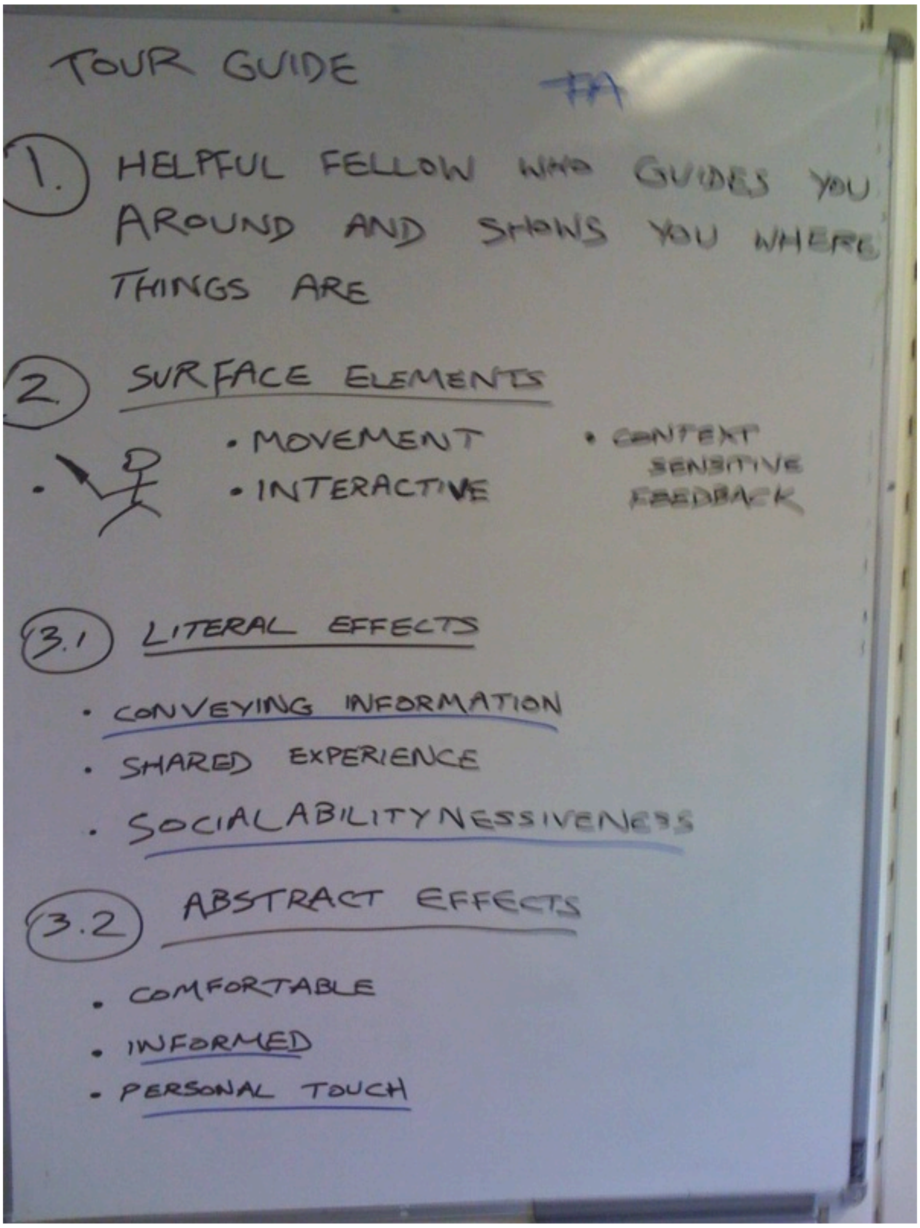


Figure 7-1 Teasing Apart tour guides

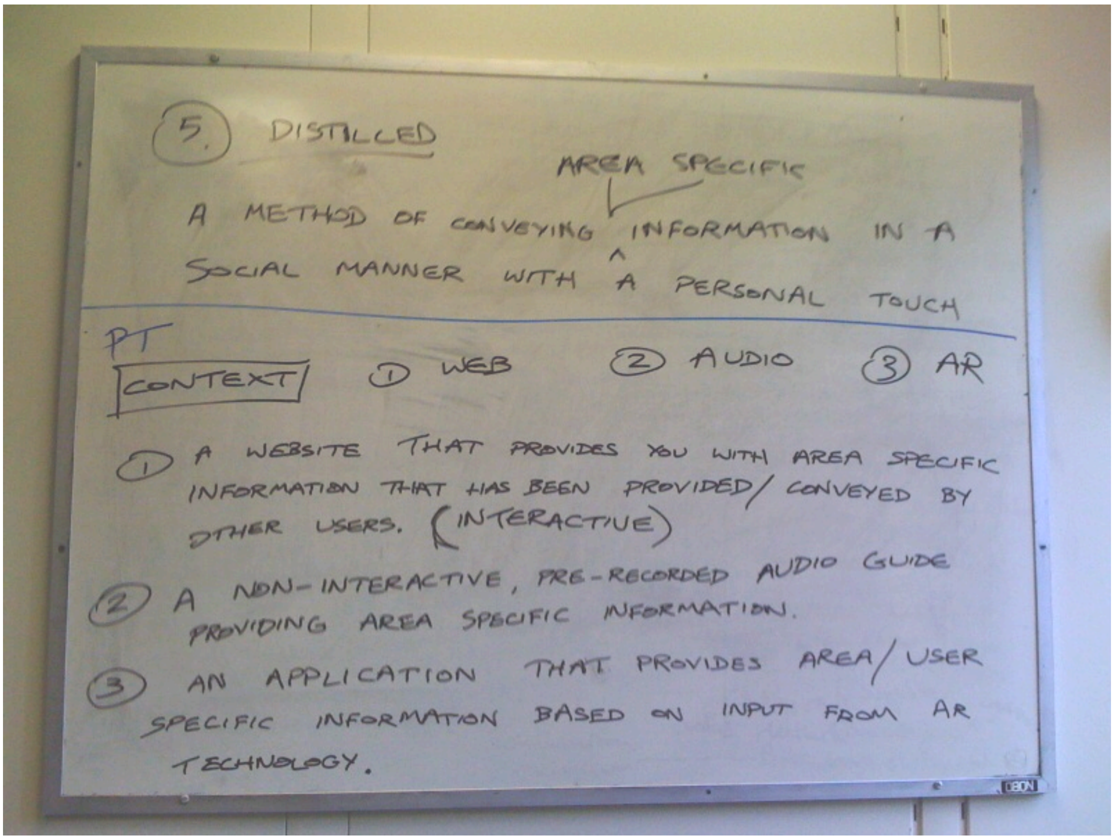


Figure 7-2 Piecing Together tour guides

7.2.3 Study outcome

Figure 7-3 shows a slide created by the Blapr team to explain their system. Their system uses IBM messaging technology to provide information on the location and status of end users via a number of mechanisms shown on the slide. (Notes, DB2 and Sametime refer to IBM products used within IBM, and 'map tool' is an internal IBM tool. 'QR codes' are two-dimensional bar codes.)

The Blapr team felt their project largely ran as they expected, although several members of the team commented about over-estimating what they could achieve in the time they had. Participant D felt the team sometimes suffered from a lack of clarity about the problem they were solving.

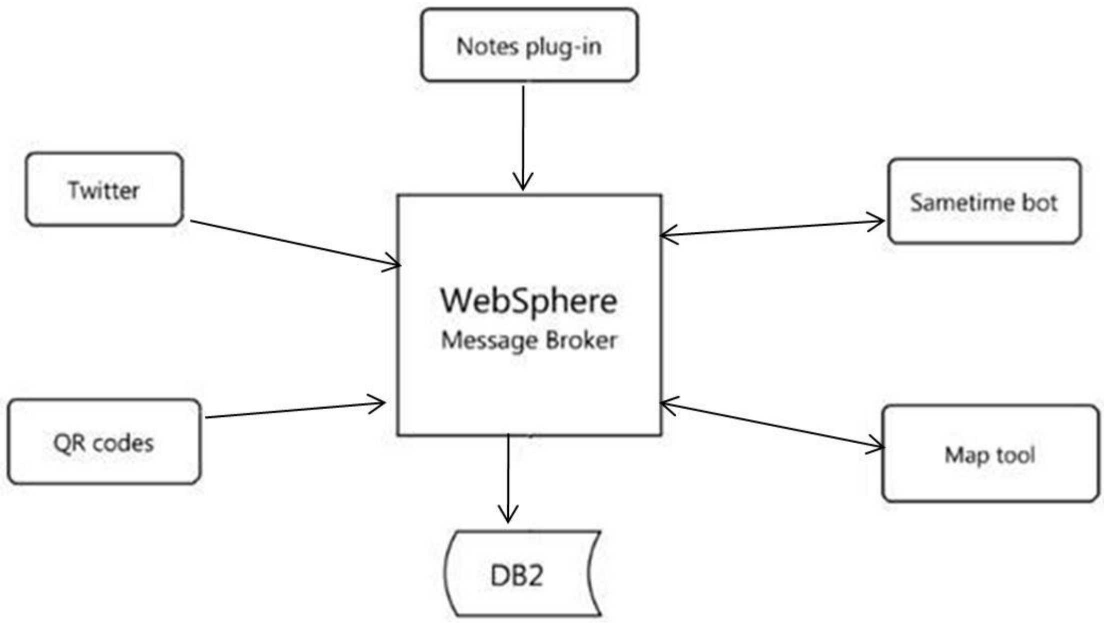


Figure 7-3 The Blapr system

7.3 Study 2: Social Networking on the Spoken Web

The next study involved using TAPT to facilitate the redesign of social networking websites for the spoken web, for use in rural India. This work was conducted as part of IBM's broader Spoken Web project[22]. Spoken Web is based on the premise that many people, particularly in developing regions, have no access to the traditional web. The Spoken Web (or World Wide Telecom Web) is a voice-driven parallel to the web, consisting of voice applications created by users and hosted in the network, accessed through phones. It aims to enable the underprivileged to create, host and share information. One facet of this work is the development of social networking facilities analogous to those found on the traditional web.

Participant F, an IBM India employee, is currently working on the social networking facet of the Spoken Web project. Participant F was asked if she would like to use TAPT to support her work, and she said yes.

[22] http://domino.research.ibm.com/comm/research_people.nsf/pages/arun_kumar.wwtw.html, last accessed November 2010.

7.3.1 Approach taken by Participant F

Participant F travelled to IBM Hursley for one day in August 2010, where in the morning she and the researcher ran a focus group with five local IBM employees. Participant F sought subjects with varying level of experience with social websites to gain insight into experiences across a number of perspectives.

Participant F first administered a brief, one-page questionnaire to the participants, intended to elicit their prior experience with social networking websites. She then ran a 45-minute open discussion session about social networking sites, covering broad areas such as what people get from social networking in general, as well as particular aspects such as events, fan pages, and mobile versions of sites. For a further 30 minutes, subjects worked together to apply the Teasing Apart phase of TAPT to two aspects of social networking: 'status updates' and 'notifications'. These two facets were chosen by Participant F as key to social networking.

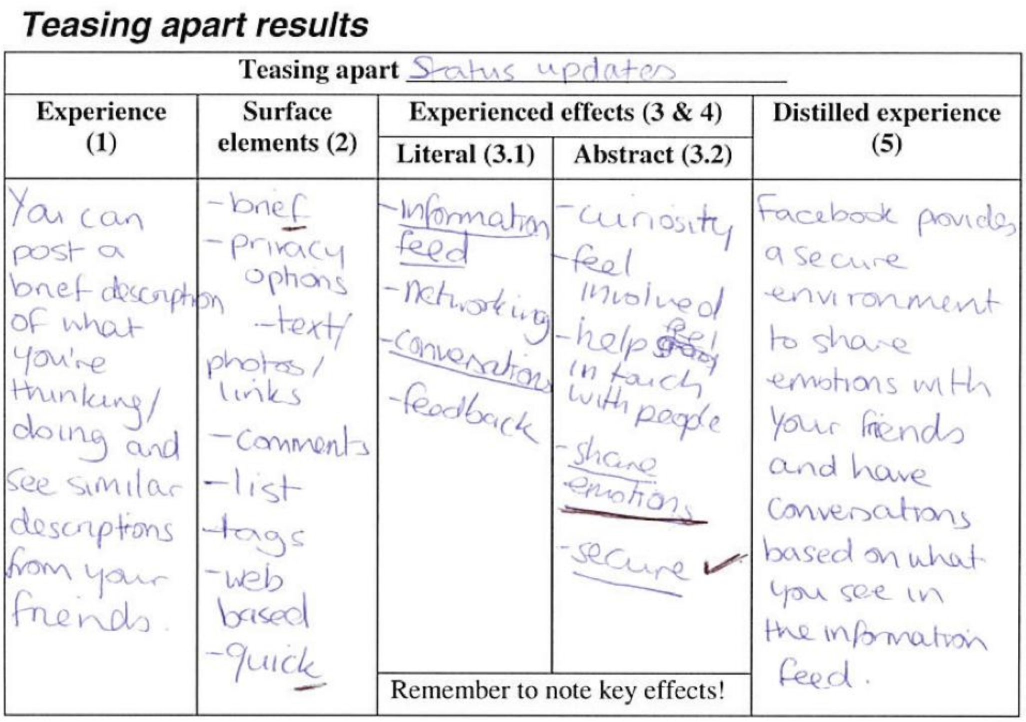
Members of the focus group were given a description of TAPT that matched that given to participants of the comparative evaluation (see Appendix E: Comparative Evaluation Materials), with one change: the Facebook-related example was removed, because it was so close to the domain of the task that it was likely to 'lead' (influence) participants in their analysis work. This left the Christmas cracker example. No substitute example was included because Participant F and the researcher would be present to provide clarification if required.

After the focus group had closed, Participant F looked over the resultant materials and went about the process of Piecing Together the two distilled experiences in the context of the social web, with support from the researcher. This took about 40 minutes. The Piecing Together table resembled that given to participants of the comparative evaluation, except that one cell was added to the table for optional diagrams and sketches. In the event, this was not used, although Participant F conducted some brainstorming on blank sheets of paper.

Participant F was interviewed three times: at the start and end of her August 2010 visit to Hursley, and three months later.

7.3.2 TAPT output

Figure 7-4 shows a Teased Apart analysis generated by the focus group. Transcriptions of the TAPT artefacts are in Appendix Q: Case Study Artefacts.



Teasing apart results

Teasing apart Status updates

Experience (1)	Surface elements (2)	Experienced effects (3 & 4): Literal (3.1)	Experienced effects (3 & 4): Abstract (3.2)	Distilled experience (5)
You can post a brief description of what you're thinking/doing and see similar descriptions from your friends.	-brief -privacy options -text/photos/links -comments -list -tags -web based -quick	-information feed -networking -conversation -feedback	-curiosity -feel involved -help feel in touch with people -share emotions -secure	Facebook provides a secure environment to share emotions with your friends and have conversations based on what you see in the information feed.
		Remember to note key effects!		

Figure 7-4 Teasing apart status updates

7.3.3 Study outcome

Participant F was interviewed in November 2010. Ideally the interview would have been postponed as Participant F's work was still ongoing at this stage, but time constraints on the part of the researcher necessitated gathering the data at this time.

Since August 2010, Participant F had built a test application and run it with 15 Indian users with low literacy and cell phone access. She found one unexpected aspect, which was that although the subjects were paid to access the site, they did not want to use it due to a lack of trust. Additionally, they found it difficult to navigate the voice site. She felt they found it difficult to understand the concept of social networking: "It's something that implicitly they do […] but when you call it social networking they don't really know what to do with it."

7.4 Study 3: Understanding Genres of Gaming

The researcher was approached by two colleagues at the University of Southampton who were conducting research in the arena of e-learning and educational gaming. They proposed the use of TAPT as a way to better understand how gaming and learning experiences work, towards improving their understanding of how to build games that provide effective learning experiences.

The motivation for this work lay in the difficulty of incorporating learning elements into games, exemplified by Kelly's description of building *Immune Attack*, a computer game to teach high school students about immunology (Kelly, 2007). The stakeholders in the design process ranged from game designers to biologists and teachers to immunologists, all of whom were pulling in different directions to make the game 'better'. Compromises had to be made in order to provide a fun and engaging game that integrated factually accurate subject matter.

The proposed methodology was to use TAPT to Tease Apart lesson plans for teaching a specific topic and to Tease Apart games from specific genres, in order to gain insights into the experiential aspects common to both teaching and gaming (Hooper, 2010a). The distilled experiences of each could be used to inform the game design process, while the insights into genres of gaming would complement ongoing research by Frazer on deconstructing features of game genres (Frazer, 2008).

7.4.1 Approach taken by Participants G and H

An initial study was run towards gaining insight into the experiences of people playing games within five specific game genres (identified by one of the two participants). Participant G built five lists of games (one list for each genre), selecting those that he felt represented the genres identified in his research, and that were likely to have been played by participants.

The researchers elicited TAPT analyses of games by placing the following items in a communal coffee room at the University of Southampton for one week:

1) Instructions for participants
2) A list of five games to be analysed (changed daily)
3) Pens and forms for TAPT analyses
4) A box for depositing completed forms

Members of the relevant building were emailed with a brief description of the available resources, and invited to analyse the experience of playing any of the listed games when they were in the coffee room. The analyses were conducted at subjects' leisure: Participant G made this decision to leave forms in the coffee room, aiming to recruit "People who'll take something and run with it rather than sitting down formally." The study ran for one week.

7.4.2 *Pilot study*

A pilot study was conducted in June 2010 to refine the experiment materials. Before the study, Participant G modified the TAPT instructions to make them relevant to the gaming domain, particularly by including an example grounded in the domain and Participant H reworked the TAPT form to make it fun, informal and collaborative. Figure 7-5 shows this reworked version of the form. It includes various shapes and colour, in addition to text oriented in different directions. This re-orientation of text was intended to make the form accessible to multiple participants sitting across a table from one another.

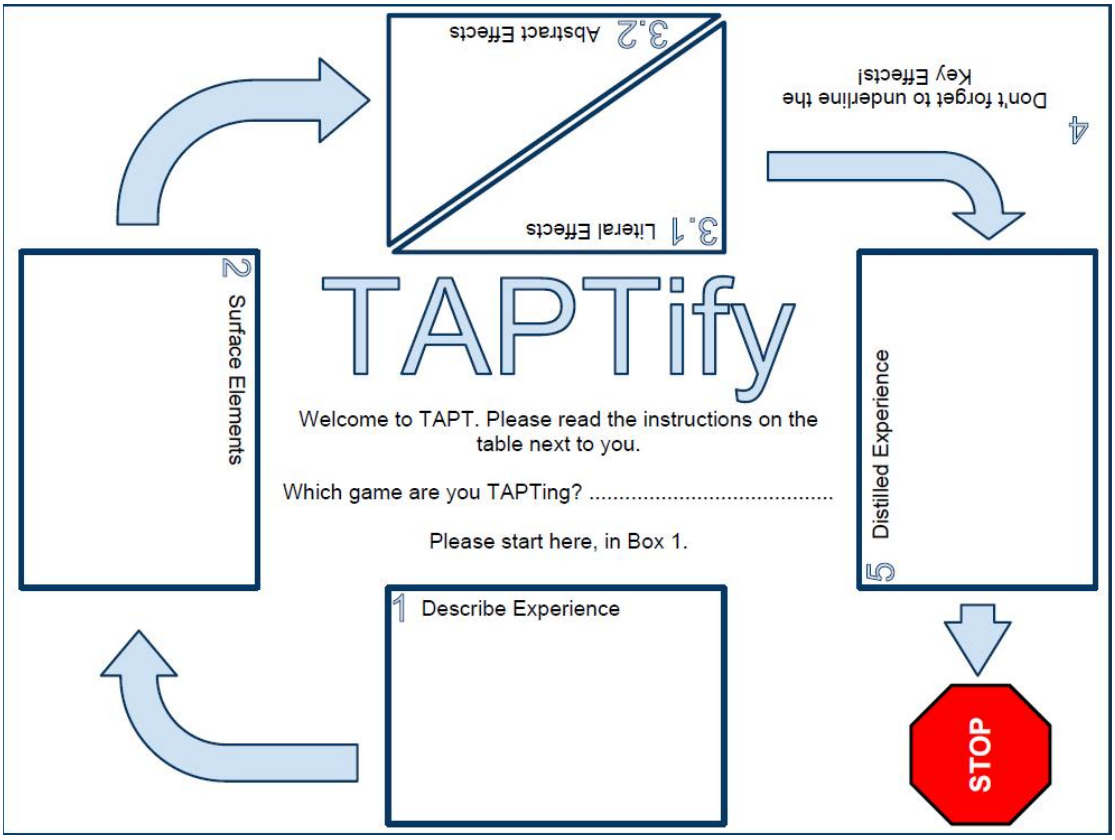


Figure 7-5 First version of TAPT form produced by Participant H

Feedback from the pilot resulted in various changes:

- Clarifying the difference between the starting description and the distilled description.
- Merging the instructions and example into a large annotated example form.
- Providing text with one alignment rather than rotating around the sheet, to support subjects reviewing lists of effects when writing the distilled experience.
- Providing the same numbering depth for all steps (making steps 3.1 and 3.2 into steps 3 and 4).

In light of this feedback, Participant H reworked the form to the version shown in Figure 7-6, which kept the flow-chart concept of the first version. In addition to this material, two copies of the form completed with handwritten examples were left for subjects' use.

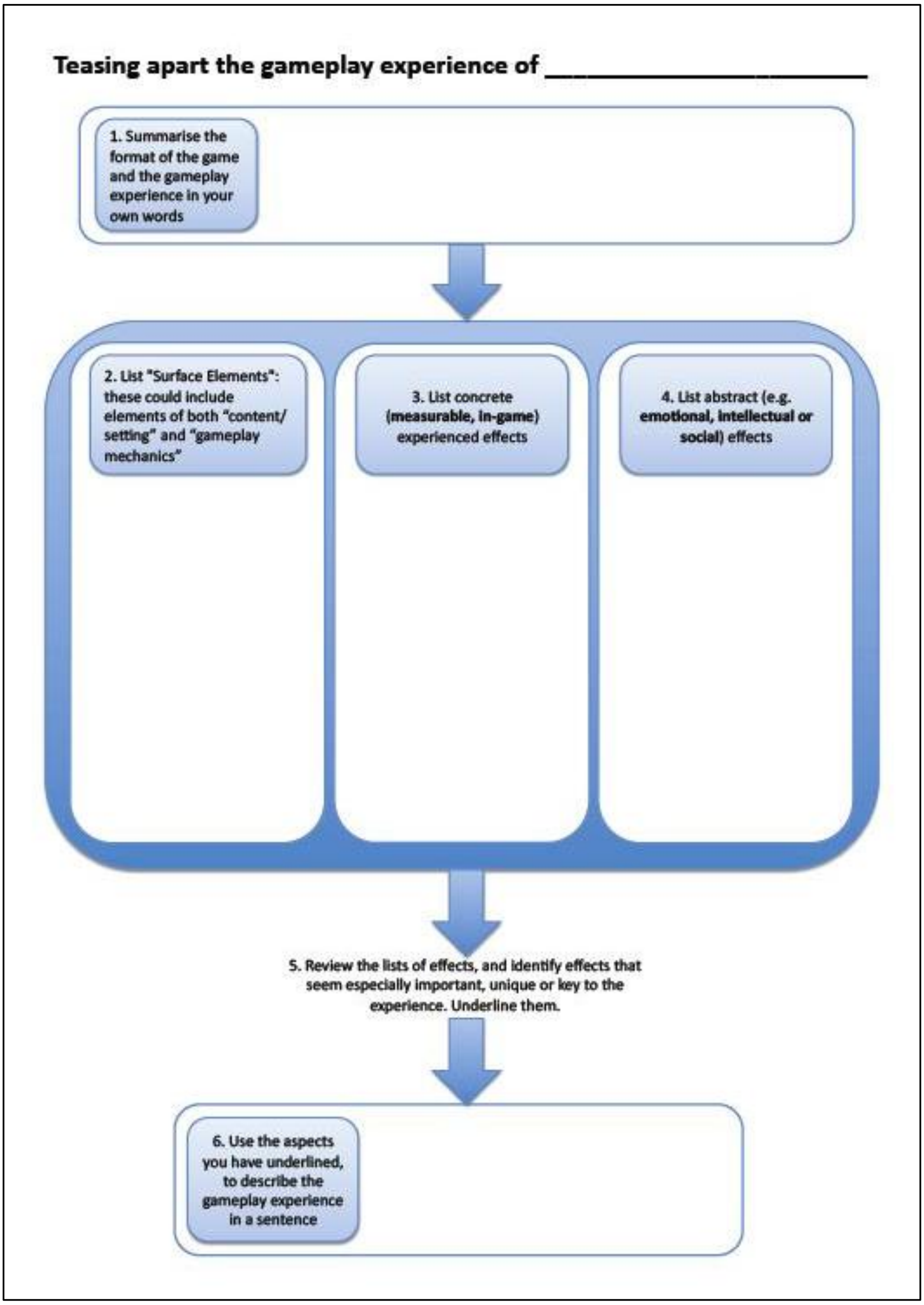


Figure 7-6 Final version of TAPT form produced by Participant H

7.4.3 TAPT output

Fifteen analyses of games were submitted over the course of the study, consisting of five analyses of puzzle games, four of strategy games, one of a role-play game, two of action games and three of board games. Figure 7-7 shows an example analysis.

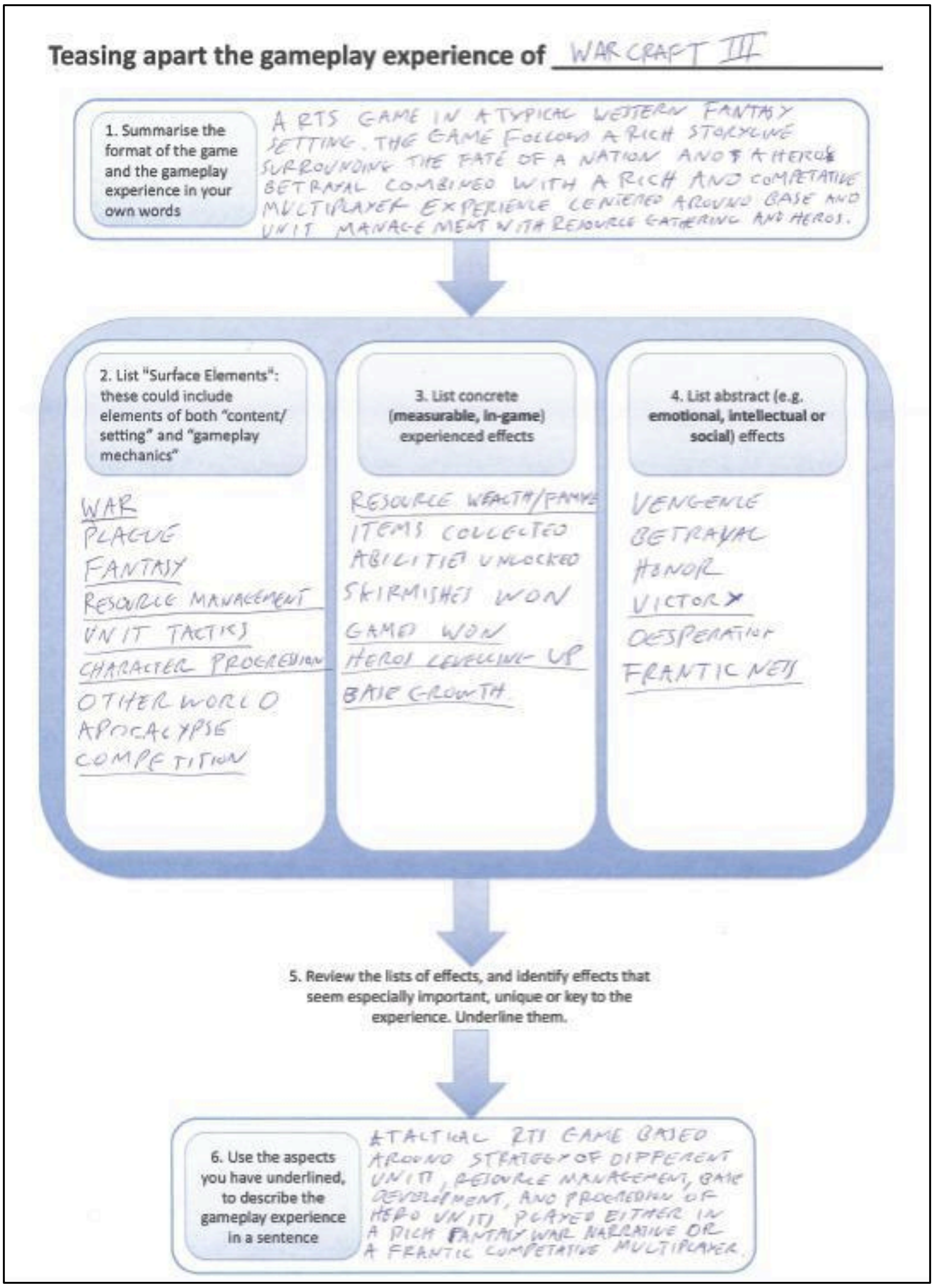


Figure 7-7 Sample TAPT artefact from gaming study

7.4.4 Study outcome

Participants G and H did not obtain as many results as they had hoped for. Participant G said “We didn’t have enough people doing it. There wasn’t enough variety in the results to make decent correlations,” and suggested this was due to “The leisurely approach we took to leaving things out, letting people do things as they wanted."

Participant G confirmed that he would run the study differently were he doing it again, primarily by being more realistic in goals of data collection. He described picking one genre and recruiting twenty or thirty participants "To try and noise cancel the differences in how people have approached the different games." He talked about the difference between checking similarities within and between genres, and approaching those two areas with two separate studies.

Participants G and H conducted some initial analysis of the data. They noted various facets of the set of analyses, including overall commonalities (for example, literal effects such as 'points', 'progress' and 'winning' and abstract effects including 'heroism', 'vengeance' and 'rebellion'), and commonalities within genres (such as 'order' as an abstract effect of puzzle games and 'politics / negotiation' as a surface element of strategy games). They were surprised to note that abstract effects did not refer a great deal to 'success' or 'failure'.

The participants noted different interpretations of what is meant by surface, concrete or abstract effects. For example, surface effects varied from broad themes ('Mystery' and 'Scientific theme') to specific elements ('Keycards', 'Crates'). They also noted that sometimes the change between the starting description and the distilled description was not as expected, with the second description becoming less abstract and longer, rather than more abstract and shorter, and sometimes there being very little difference at all.

Participant G speculated that deconstructing games may produce fewer results than deconstructing other experiences, because games themselves are in a sense already distilled: "People have looked at the real life scenario, like you're a soldier doing stuff or you're building an empire or you're racing a car round something, and they look at that and think 'right, we can't necessarily simulate this exactly, what are the important bits?' […] They'll implement that rather than a perfect simulation." He speculated that this might be why the initial and closing descriptions were similar in places, because "as an experience it's too distilled already."

The participants' primary comment was that the results were fewer in number than they had hoped, but that those results began to highlight relevant issues and were very helpful in terms of informing future work.

7.5 Study 4: Understanding Social Location-based Tools

The researcher approached Participant I, an Associate Professor at the University of Bergen who conducts research into storytelling online. Upon hearing about the TAPT process, Participant I proposed a study into the experiences of social location-based tools, specifically geocaching (a treasure hunt enabled by location-aware technology) and Gowalla (an application for sharing one's location)[23].

Participant I explained that the increasing popularity of smartphones in Norway has driven increased use of location-based services, which appeals to her research interests: "Location stuff is becoming mainstream and it's really interesting as a new hypertext thing." She said she was driven by a fascination with the area, saying "This is cool and new but what's it for, and why?"

7.5.1 Approach taken by Participant I

The researcher ran a small pilot study which involved asking a PhD student at Participant I's university to apply the Teasing Apart phase of TAPT to a digital experience of her choice (reading Facebook). This was to ascertain instruction clarity and how much time to allow subjects: although TAPT had been used many times previously, this was its first use by people from a non-engineering background and from a non-British culture. The pilot led to two minor changes being made to the TAPT instructions for clarity: explicitly stating which lists were to be checked for key effects (the lists of literal and abstract effects), and emphasising that there is no wrong answer when Teasing an experience Apart.

The main study consisted of two one-hour focus groups run by Participant I. The groups consisted of enthusiastic and experienced users of each of the two applications. Subjects were asked to collaboratively apply the Teasing Apart phase of TAPT to the experience of using geocaching or Gowalla.

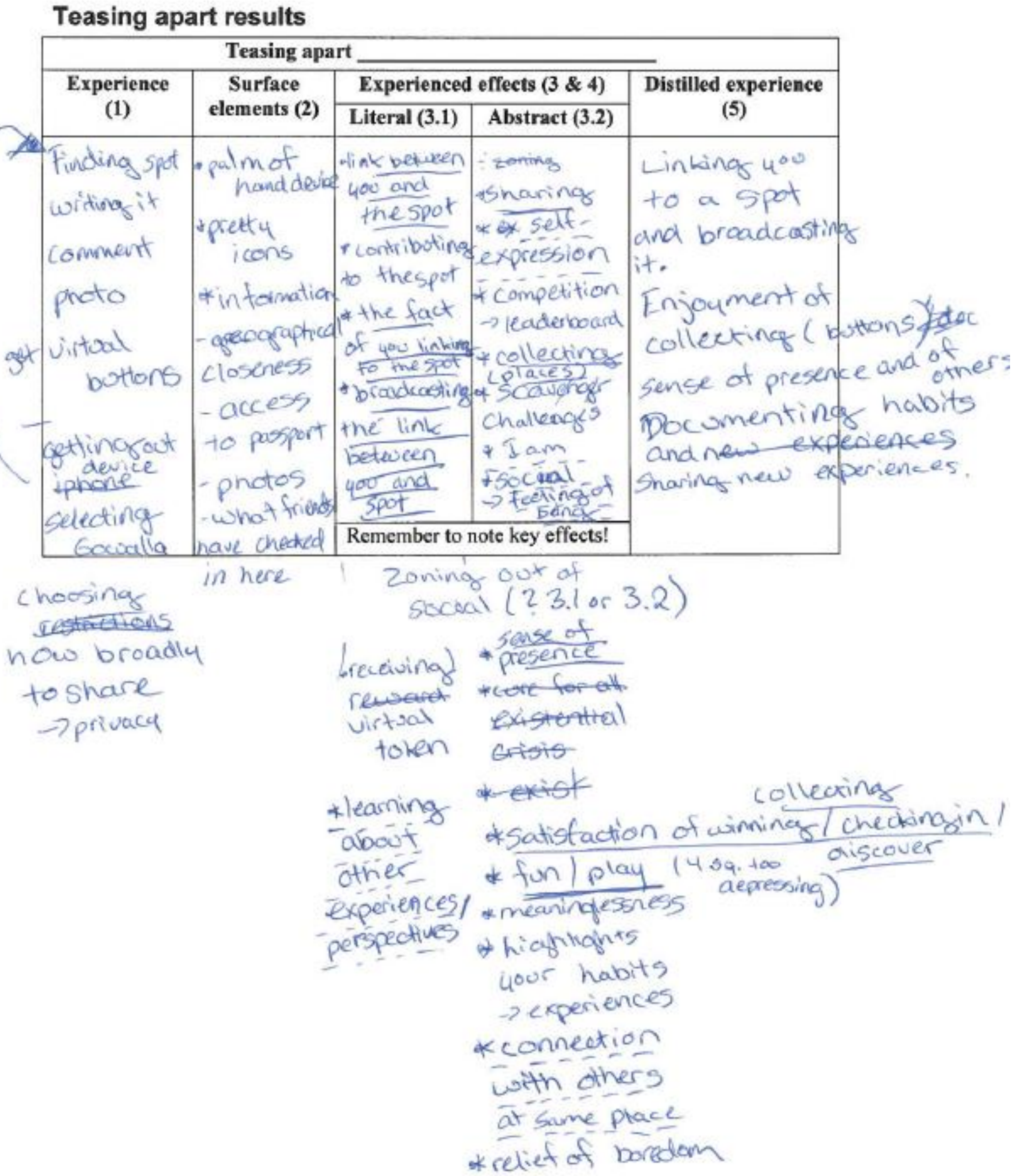
The resultant TAPT analyses were then analysed by Participant I and the researcher. This process involved: taking keywords from the analyses as a springboard; exploring similarities and differences between the two analyses; considering the results from a hypertext perspective by identifying links and nodes

[23] http://www.geocaching.com/ and http://gowalla.com/, last accessed December 2010.

and considering how the results link with theory such as patterns of hypertext (Bernstein, 1998).

7.5.2 TAPT output

The focus groups applied TAPT to geocaching and Gowalla. Figure 7-8 shows the analysis produced by the Gowalla group. Both artefacts can be seen in Appendix Q: Case Study Artefacts.



Teasing apart results

Teasing apart				
Experience (1)	Surface elements (2)	Experienced effects (3 & 4) Literal (3.1)	Abstract (3.2)	Distilled experience (5)
Finding spot writing it comment photo virtual buttons getting out device ~~iphone~~ selecting Gowalla	• palm of hand device *pretty icons *information - geographical closeness - access to passport - photos - what friends have checked in here	- link between you and the spot * contributing to the spot * the fact of you linking to the spot * broadcasting the link between you and spot	: zoning *sharing * ~~ex~~ self-expression * competition → leaderboard * collecting (places) * scavenger challenges * I am * social → feeling of being	Linking you to a spot and broadcasting it. Enjoyment of collecting (buttons, ~~etc~~) sense of presence and of others Documenting habits and ~~new experiences~~ Sharing new experiences.
		Remember to note key effects!		

Figure 7-8 TAPT applied to Gowalla

7.5.3 *Study outcome*

Participant I noted that there were two sets of outcomes from the study: outcomes about the method and outcomes about the experiences that she asked the subjects to describe.

Regarding the method, she commented that:

- The two focus groups responded to TAPT in very different ways[24].
- It was useful to acquire key words that users agreed upon and that could be used as a springboard for further analysis ("It was very useful getting key words that users agreed upon and using them as a springboard to find the connections to do more analysis […] a very good result there").
- It was interesting to observe the emerging process, such as the post-focus group work relating TAPT results to existing theory ("The way [the TAPT analyses] set us onto thinking further in terms of relating things to theory, that's another result").

Regarding results, she observed:

- The Gowalla description sounded like an advert for Gowalla. There was a comment from a subject that they were enthusiasts and therefore they were trying to make a description which defended that interest.
- An insight that Gowalla emphasised 'being' and geocaching emphasised 'doing'.
- Links between the results and existing hypertext theory. For example, the work revealed real-world analogies to theoretical constructs such as cycles (Bernstein, 1998) and guard fields (Joyce, 1991).

7.6 Participant Motivations and Expectations for these Studies

Before each of the four case studies ran, participants were asked to identify their motivations for using TAPT and their expectations of it. At the close of the case studies, they were asked whether TAPT had met those expectations.

---

[24] The larger group was dominated by one subject, who was in charge of writing. The second group, which only consisted of two participants, approached the task differently: both members applied TAPT separately before collaborating to produce one agreed-upon analysis.

7.6.1 *Motivations*

The Blapr team was asked to use TAPT, but other case study participants chose to use it for various reasons:

Participant F (Spoken Web) explained how her work involves re-providing experiences analogous to those of people using web-based social networking via voice interfaces, and explained that a study using TAPT seemed suited to this work and “A good starting point” for understanding online social experiences.

Participant G (Gaming) talked about his desire to deconstruct gaming and learning experiences towards being able to facilitate the building of games to support learning. He said he did not consider any methods apart from TAPT, which “made a lot of sense” to use because he wanted to do deconstructive work and understand the elements which comprise games and which of those are most important. Participant G also commented on an interest in other people deconstructing games because he wanted their perspectives as potential end users: “See if they can actually make sense of [my work on identifying genres of gaming]. See if they can do practical things with it. It’s no good if it’s only useful to me, it’s got to be useful in and by the field.”

Participant H (Gaming) said TAPT “Seemed appropriate for getting people thinking about how to deconstruct a certain activity or experience, which is exactly what we needed to work out what makes up a game.”

Participant I (Location-based Tools) remarked on TAPT’s promise of “A description of an experience that's potentially technology-neutral.” She found this “very, very interesting” especially as her interests were in the social and experiential aspects of the technologies in question.

In summary, participants had various motivations for using TAPT: they wanted to redesign experiences in new contexts, get subjects to deconstruct experiences and understand experiences in a technology-neutral fashion.

7.6.2 *Expectations*

When interviewed at the beginning of their work, **Blapr** participants identified two specific hopes and expectations.

- They felt TAPT would have a positive impact on their ongoing work, helping them get “More ideas on specific details” of experiences that they wanted to reproduce (Participant A)
- Participant E felt it would enable them be more original: “It means we take basic concepts, rather than just copying something, we maybe could use them in a better way or a different way. I think that's an advantage of using [TAPT].”

After their work was completed, the Blapr team seemed to feel that TAPT had helped them drill into the experiences they wished to reproduce. For example, Participant D talked about TAPT’s efficacy at helping them include “subtle, implicit” aspects (see Table 7-3).

Regarding originality, four of the five Blapr participants felt TAPT helped them design things in an original way rather than doing a straight replication of existing functionality. Comments included “It did allow us to take the best part of those things” (Participant E) and “I think we arrived at [our design] in a very unique way” (Participant D). Participant B felt TAPT did not help the team design things in an original way because TAPT was about replicating experiences, concluding that because of this “The ideas generated from it aren't as novel if you were doing, say a brainstorm.”

Participant F also identified goals for her **Spoken Web** work:

- Achieving “At least some sense of what people enjoy about social networking.”
- Getting enough data that she would not have to run a user study in addition to her study using TAPT.
- Using data from TAPT to drive design decisions when building the spoken website (“Hopefully if I can distil down that top experience enough I can design a user application in the voice domain.”)
- Using the identified abstract effects (which she felt were of particular value) to model the experience identified by participants.

After the study, Participant F was asked whether these hopes were met. Excepting the final goal, she said that they were: she said she had “definitely” acquired information on what people get out of social networking. She did not need to conduct a user study after the TAPT study, saying “I got sufficient data from the study [with TAPT].” Participant F was “definitely” able to use data from TAPT to design her application (“I have used the feedback I got through TAPT. It affected [features such as the implementation of message notifications]”). Of modelling abstract effects, she reported she had not yet come to this part of the work, although she still planned to conduct the work.

Participant G identified the following hopes for the **Gaming** study:

- Taking a "practically useful" step towards designing educational game experiences by finding out if "A person can deconstruct the game properly, and once we've done that see if it can be built into something useful."
- Gaining unexpected insights from the TAPT analyses in their own right ("Seeing what words people choose to describe these things […] there's freedom in those instructions to express things how they want, so if things come out of this that I haven't considered before then that's good").

Participant H also identified some goals:

- He hoped to "Identify the main experiences of a game, maybe summarise the ones common to all games. That's why we split it into genres, as games in each genre are more likely to have shared elements."
- "Create a set of expected elements in each genre of games" which could inform his future research in gaming, towards choosing what type of game might be most appropriate for various learning activities.

Participants G and H felt that the study ran as expected, but finished with fewer results than they had hoped for, or than were useful for conducting a significant analysis into the elements of games or genres of game. Participant G noted that there were initial insights into some differences and similarities within genres, but that a bigger study might reveal more: he felt the output had the capacity to be of use in designing game experiences, and identify the main experiences or elements involved in games, if the results were greater in number such that statistical significance could be inferred.

Both Participants G and H felt they gained insight into the perspectives of the subjects in terms of the ways in which they chose to deconstruct games (which were oriented around either mechanisms or themes), and Participant H noted an interest in seeing what subjects saw as key to a game and why they enjoyed a particular genre more than another because of those features of interest to them.

Participant G felt that he could begin to see some "nice" insights: when asked to elaborate, he talked about seeing trends across genres such as action and strategy games, noting that "We can't necessarily call that scientific victory with three or four results per [genre], but it shows there is a capacity for things to match up once you normalise the language […] there’s definitely mechanical things that people have looked at here that could be useful in designs in future.”

Participant H felt the study was useful for learning about the process and how to conduct similar investigations in the future. He also felt he had learned something about what people think make up some of the games, which could be the first step towards building an ontology of game elements.

Participant I identified her hopes for the study into **Location-based Tools**.

- That TAPT would let users express their experiences: “This is [conducted] by the users […] perhaps I have prejudices that I'm not even aware of […] this method will maybe allow me to cut straight through [those prejudices].” She later said “It's different asking users to do that [analysis]. I'm interested.”
- To understand what Gowalla and geocaching are about in a deeper sense. Of Gowalla she said, “Yes, it's a system for checking in and telling people you're at such-and-such a place, but I'm hoping this might get beneath that, maybe there's something more fundamental.”

After the study, Participant I said she had “definitely” gained insights into the perspectives of participants and how geocaching worked on a superficial level, commenting “The emphasis they placed on the secretiveness and the playacting, I had no idea, and I thought I had a reasonable idea of what geocaching was.” She added “That wasn’t in their initial description, that was stuff that came out through this process.” Participant I’s understanding of Gowalla in terms of superficial aspects did not change: “I think it’s more about what was emphasised. I didn’t learn anything new as such, as I know that service better, but definitely the emphasis and the way it was discussed was very useful.”

When asked if she felt she now had insights into what Gowalla and geocaching are on “a deeper level” she said “Absolutely! I think probably even more than I’d imagined.” She said that she felt the subsequent analysis of the output of each focus group was where she really gained that understanding, adding “But that’s probably because we’re able to compare, they only had one to discuss, we’re looking at it from above.”

Table 7-1 summarises participants’ goals and whether these were met.

Case study	Goal	Met
Blapr	Get more ideas on specific details of experiences they want to reproduce	Yes
	Be more original by using basic concepts in a different way	Yes (mostly)
Spoken Web	Understanding what people get out of social networking	Yes
	Getting sufficient data from one study	Yes
	Getting data to drive design decisions	Yes
	Using abstract effects	N/A
Gaming	Progress towards designing educational game experiences	Yes
	Insights from analyses in their own right (into users, games, genres)	Yes
	Identify experiences of games	No
	Identify experiences of genres	No
Location-based tools	Users express their experiences	Yes
	Deeper understanding of Gowalla and geocaching	Yes

Table 7-1 Summary of hopes and whether they were met

TAPT largely fulfilled participants' goals, with the following exceptions:

- Participant E of the Blapr team did not feel it prompted more original designs (although Participants A – D felt that it did)
- Participant F has not yet tried to use the abstract effects from TAPT in the Spoken Web study
- The gaming study yielded insufficient results to gain insight into experiences of games and genres, although some initial insights were noted. The participants appeared to feel that TAPT would be able to fulfil their goals if they had more data
- Participant I did not feel she gained insight into what made the services she was interested in so compelling, although she appeared to feel with hindsight that perhaps the goal was unrealistic.

TAPT's general ability to fulfil participants' goals suggests that participants were realistic in their expectations of the method and that TAPT is able to support goals such as understanding specific and deep aspects of experiences (with an emphasis on users providing this material in their own words) and building original and informed designs.

7.7 How TAPT was Used by Participants

This section discusses the way in which TAPT was used by participants and their subjects, both in terms of logistics (i.e. format in which TAPT was used, such as with a focus group or group use with a whiteboard) and in terms of the focus of TAPT analyses (for example, the choice of experiences to deconstruct and the aspects of experiences which were focused upon in a deconstruction).

7.7.1 *By participants*

TAPT was used in different ways across the case studies:

The **Blapr** team used TAPT to analyse experiences similar to the experiences they were interested in, and worked as a group using a whiteboard. Participant D also applied TAPT in an evaluative sense, using the tool to Tease Apart the Blapr system based on conversations with team-mates and his own experiences. He commented that it was a "completely informal" use of TAPT.

The Blapr team reported finding Teasing Apart easier with real objects (such as paper maps as opposed to the web-based Google maps) when trying to translate experiences to a web-based modality, because use of physical objects led to a stronger context switch. Participant C reported that they chose to Tease Apart Google maps as they knew their primary goal was about mapping and routing, but that Google maps was too hard to Piece Together because the end context was too similar to the start context. Participant A said: "When we were Tearing it Apart [sic] […] we knew we wanted to Piece it Together as software, that was influencing the way we were Tearing it Apart [sic] […] we were repeating ourselves at each individual step of the experience." Participant C commented that "A lot of the experience associated with [the start domain] is the same." In other words, the team found it easier to avoid influence from the end context when the starting context was different – in this case, not software-based. The Blapr team found using TAPT for web-based artefacts prompted lots of different ideas and "Stimulated their thought processes" and stopped them from "Just building a Google maps client for the project" (Participant D).

Participant F (**Spoken Web**) ran a focus group which conducted an open discussion and then Teased Apart facets of social networking, before she used those results to Piece the facets Together in the spoken web context. She thought the open discussion before applying TAPT helped prepare the participants for using TAPT, "Because I think once they think out loud a lot of experiences, jotting them down becomes easy."

Participants G and H (**Gaming**) made Teasing Apart instructions available to users of a coffee room at their university, and collected anonymous analyses, which they intended to analyse.

Participant I (**Location-based Tools**) asked focus groups to Tease Apart specific experiences and then analysed that data herself.

Participants used results in various ways: the Blapr and Spoken Web participants used TAPT output to drive their designs, while the Gaming and Location-based Tools participants conducted their own analyses of their subjects' TAPT output.

The Blapr team used TAPT to hone their project direction, illustrated by various comments:

- "We've now got a list of things we want to aim for." (B)
- "At the beginning we didn't really know which way we were going, what features we were going to have […] TAPT helped us decide which way, what functions." (C)
- "[TAPT provided] "an easy way to highlight key aspects which need to be captured […] and to ignore those aspects which are less important." (D)

The Gaming and Location-based Tools participants both began to relate TAPT output to existing theory, with Participant G (Gaming) beginning to discuss exploration of (game) space as linked to constructivist and experimental learning and the goal-oriented nature of action games (saying "Even at this initial level it's quite cool"), and Participant I (Location-based Tools) talking about linking the TAPT output with hypertext theory, particularly patterns of hypertext: "The way that set us onto thinking further in terms of relating things to theory, that's another result."

In summary, participants used TAPT to understand similar experiences to the experience for which they were designing (Blapr); experiences that they wished to redesign for new contexts (Spoken Web, Gaming); and experiences that they wished to understand in their own right (Location-based Tools). Participants used TAPT collaboratively around a whiteboard (Blapr); via focus groups (Spoken Web, Location-based Tools); and via anonymous participants (Gaming).

Various other findings came to light:

- Participants reported finding TAPT easier when applied to experiences in a different modality to their end context. This has implications for the type of situation to which TAPT is most appropriate, suggesting it is most suited to supporting the translation of experiences across *different* modalities.
- The Blapr team appeared to go through a process where they learned which experiences were more and less helpful to Tease Apart, providing further evidence that there is a learning curve with the process.
- The Blapr team used TAPT in an unexpected way, to analyse experiences similar (but not identical) to the experience they were interested in. This represents a new way to use TAPT.
- Participant A (Blapr) reported conducting Piecing Together in a fashion that was led by design ideas he already had. It seems likely that any practitioner conducting design will be led by their existing ideas, but this comment points to an avenue of future work examining the extent to which design methods impact pre-existing ideas (and vice versa).
- Participant D talked about the use of TAPT to build products which offer *familiar* experiences, not necessarily to build products to replicate experience: "The use of TAPT at the start of a project is best felt when deciding the most appropriate format to deliver the solution based on users' previous experience and expectations for the presentation of information."

7.7.2 How TAPT was used by subjects

Participant G (gaming) felt that people deconstructed games as expected, but noted that two types of deconstruction happened: "People sometimes […] deconstructed game play mechanisms more explicitly than they have the theme of the game." He felt that both were valid approaches, but that in such a small set of results it was difficult to draw correlations between them.

Participant G found the study useful for seeing how the experiment worked, and how differently subjects chose to describe the deconstructed games and experiences. "The main thing is it helped us understand, when left to their own devices, how people choose to deconstruct these experiences."

On subjects deconstructing game play mechanisms or themes and stories within games, he said "Maybe we should have been more specific about which part we wanted people to deconstruct, or presented games where one of those wasn't as strong an element."

Participant H (Gaming) felt the games analyses were not as similar as he expected, saying “Even between the same games people came up with different things. I guess we should have realised this before, it’s quite a subjective process. I hoped the process would allow people to come up with the same things, or the people we were asking would have the same views about the games and what the important parts of them are.”

Participant H said that the results were too few in number to narrow down game genres to link them with educational aspects, and that the mixture of theme and game play elements confused matters further: “People pick out one or two of the game play elements and the rest are ‘oh it’s about war or it’s about this’ and that confuses the data.”

He noted that the results highlighted some but not all of the experiences in each game, returning to a theme of subjectivity. He talked about applying more iterations, or subjects working separately then collaborating to join ideas on the same game towards achieving “A more holistic, full analysis of the features.”

Participant I (Location-based Tools) felt there was “a lot of value” in having groups interact and come to a shared result, explaining that “In both the groups there were certainly things that came out through discussions that the individuals might not have put down at the start.”

Participant I found that TAPT was presented to each of the focus groups in the same fashion, but the two groups responded in very different ways: the first group was dominated by one subject who was in charge of writing while two or three group members seemed disengaged through the second half of the session. The second group, which was smaller and only consisted of two subjects, approached the task differently, with both members applying TAPT separately before later collaborating to produce one agreed-upon analysis. Participant I was surprised that the first group did not share the workload more evenly.

Participant I noted that groups did not work as expected, for example with some members of the Gowalla group being quiet: “I thought the Gowalla people would be much more vocal […] that surprised me.”

In summary, comments about subjects’ use of TAPT concerned unexpected approaches to deconstruction (in terms of which facets of experience were focused upon and in terms of how TAPT was applied) and the recruitment of many subjects to overcome individual subjectivity.

7.7.3 Time required to learn and apply TAPT

Where possible, a note was made of the time participants or their subjects spent applying TAPT. Data were as follows:

Case study	TAPT user(s)	Task	Time taken
Blapr	Participants A - E	Tease Apart and try to Piece Together Google maps (first usage)	60 - 90 minutes
	Participants A - E	Tease Apart and Piece Together signposts and tour guides (subsequent usage)	45 - 60 minutes for each experience
Spoken Web	Focus group of five subjects	Tease Apart status updates then notifications	30 minutes for both experiences
	Participant F	Piece Together status updates and notifications	40 minutes for both experiences
Gaming	Anonymous subjects	Tease Apart various games	Unknown (subjects were not observed)
Location-based Tools	Focus group of five subjects	Tease Apart Gowalla	60 minutes
	Focus group of two subjects	Tease Apart geocaching	45 minutes

Table 7-2 Time taken to apply TAPT in case studies

The extent to which Blapr speeded up in using TAPT after their first effort is notable. Participant C (Blapr) reported they were "quite confused" at the beginning of the TAPT process, specifically in terms of understanding the difference between literal and abstract effects, while Participant A (Blapr) found it "much more useful" on the second usage, "Once we'd got idea of how to do it from the first time". It is possible that they worked more quickly when analysing signposts and tour guides as these were a more appropriate choice of experience, but nonetheless this greater speed and comfort applying TAPT a second and third time strongly implies the existence of a learning curve (discussed in Section 5.4.8).

The Spoken Web participant and subjects were particularly fast to apply TAPT (taking 65 – 70 minutes to Tease Apart and Piece Together two experiences). The researcher was present during this work and provided help where required, which probably contributed to their relative speed.

The two focus groups to Tease Apart location-based tools worked at different speeds. It is possible that the larger group took longer as they engaged in more discussion and negotiation of the elements and effects of using Gowalla, while the pair who looked at geocaching largely agreed with one another from the outset.

Blapr Participant B reported that the longest task when using TAPT was "Making lists of what the experience is composed of -- especially the emotional aspects": this returns to comments from Section 7.4.3 about perceived difficulties experienced by Software Engineers as they attempted to understand emotional facets of experience.

7.8 Comments on TAPT

In addition to gaining insights into how TAPT was used, the case studies provided data on the ways in which TAPT was or was not useful. This section presents those comments, initially discussing TAPT's fit within design, research and evaluation processes before discussing its support of understanding, replication and evaluation. Next, other identified features are listed, followed by comments on TAPT in contrast to alternative methods. This section closes with observations and comments on different ways to use TAPT.

7.8.1 TAPT's fit in the design, research and evaluation process

All teams reported that TAPT sat naturally within the processes they applied:

- The Blapr team reported that their use of TAPT sat naturally within the design process. Exceptions were Participant D, who felt their failure to define a problem made the fit less good and Participant B, who felt the project did not involve recreating an experience.
- Participant E (Blapr) felt TAPT was suitable as an evaluative tool.
- Participant F (Spoken Web) felt that her use of TAPT sat naturally within the design process and that her use of TAPT at that stage in the design process seemed "really appropriate" and "worked".
- Both Participants G and H (Gaming) felt TAPT was a natural tool for the research. Participant G remarked "I'd got ideas from the genre work and it let us implement that, I think it sat there quite well" and Participant H said "It was definitely the right time to try and break something down."
- Participant I (Location-based Tools) felt TAPT sat naturally within the research process, and was a suitable tool for the task in question.

7.8.2 Understanding experience

Many participants commented on TAPT aiding their understanding of experiences, in terms of…

Understanding details:

- "[We] identified things you might not otherwise think about, like personalisation with a tour guide. Drilling into tiny bits very useful." (A, Blapr)
- "I never thought about Google maps in this way [… TAPT] made the product more clear." (C, Blapr)
- "I think it has the potential to pick out more." (H, Gaming)

Gaining insights different to those which traditional methods would yield:

- "It made us think about how people use things in a slightly different way, especially the way it's designed using things that aren't necessarily software and bringing them into software. Normally with design, people get focused on thinking of other software and how it's already doing other things whereas this way you might think of a new method or something that's not done with software that you create in a different way. Using new methods and approaches to ideas." (A, Blapr)
- "You've just grown up with [signposts], you don't think about how is it actually being used by different people, what are their experiences of using it" (D, Blapr)
- "There is a capacity for things to match up [for comparative analysis]" (G, Gaming)

Understanding abstract aspects:

- "Personally I find it very difficult to make a connection between abstract and tangible things, how can you link those together. TAPT kind of forces you to and forces you to think about and have an actual framework to put those thoughts into. It helped me organise things in a way that I could see those links more clearly." (D, Blapr)
- "I think the most important thing for me was seeing […] the abstract experience […] looking at the activity on Facebook it's easy to come up with literal terms, but abstract terms, they are harder to come by." (F, Social Web)

Gaining rich, qualitative data:

- "Without TAPT the design process would have still gone forwards but it would have been less informed […] it's a more qualitative way of understanding things." (F, Social Web)
- "I expected there to be a maturing or shift in focus from [subjects'] first descriptions to their last descriptions, and that was interesting to see and that certainly happened." (I, Location-based Tools)
- "Well right now I'm just totally psyched to try it again! I thought it was a really useful way of generating material about, a rich description of, a technological experience." (I, Location-based Tools)

Not all Blapr participants were positive: Participant E was not sure if it was useful because "They were kind of basic things [which we analysed] […] everyone knows what a signpost is."

Some participants went into more depth. For example, Participant F (Spoken Web) said TAPT improved her understanding of social networking websites in a way that impacted her designs. She discussed how users on web-based social networks can have hundreds of friends but that the study revealed that users appeared to only keep in touch with ten or fifteen friends, which impacted the design.

Other comments about increased understanding related specifically to emotional aspects of experiences. For example, Participant D (Blapr) felt that "It's very easy in the kind of project that we're doing, where we're potential end users, to only consider our own needs and not think about other users." He added that looking at experiences such as those of signposts and tour guides made them consider feelings: "When you see a signpost it can make you happy because you're almost there, or 'ah no I have that far to go' or 'I was wrong, maybe I was going that way, I should go this way' - thinking about the feelings you get when you go through it. It's strange to think that'd make a difference, but it did."

Similarly, Participant I (Location-based Tools) was enthused by TAPT's orientation around emotions, saying "I loved that it highlights the experience of feelings attached to it because most methods don't."

Participant D (Blapr) highlighted that he felt TAPT was particularly good for understanding subtle, implicit aspects: "a signpost […] conveys not only the place it points and the distances, it also includes the direction [with its] shape and the way it's facing, so it's not explicitly written on the sign that you need to go north-north-east, it's 'this way', so it's those subtle, implicit things we were trying to… I wouldn't say we were definitely trying to include them but they have been included."

In summary, the majority of participants commented that TAPT aided their understanding of experiences in various ways (understanding details; gaining insights different from those gained from traditional methods; understanding abstract effects; gaining rich, qualitative data), and some participants commented on their increased understanding of implicit and emotional facets.

7.8.3 Replicating experience

Blapr was the only case study to engage in replication of experiences: although the Spoken Web study was about replication, at the time of the closing interview the work was still ongoing. The other two studies (gaming and location-based tools) concerned using TAPT to analyse and understand experiences, not replicate them.

Blapr participants were positive about TAPT as a tool for replication:

- "Once we had identified what we wanted to replicate [TAPT] became very relevant." "Where [replication] is what the end user is after then [TAPT] can be incredibly useful." (Participant B)
- "It did allow us to take the best part of those [other experiences]." (Participant E)

The Blapr team felt TAPT was useful for achieving their goals. For example, Participant D felt that without TAPT they might have built a system that was similar but "maybe wouldn't have achieved the goals that it has as well as it has."

All Blapr participants felt that the TAPT analyses produced at the start of the project drove their resultant designs and four of the five felt it helped them decide and focus on the important aspects of their work.

The team found TAPT useful for directing their work, in terms of…

Identifying and prioritising aspects of experiences:

- "It singled out the features of Google maps or signposts or whatever, what was most important, so they were the priority items." (A)
- "There's a section that you can put down what you think is important and I think that helped to narrow down important features to implement." (C)
- "Trying to pick different things, saying 'that's good about that, maybe we should include that', 'this isn't so good, let's not include it.'" (E)

Providing direction and structure:

- “If we hadn’t used TAPT we might have floundered around more in the design. Because it was a structured process it was quite a good way of going through things and you come up with a set of data that you’d be able to use to guide you with your design.” (A)
- “We kind of used TAPT to set the direction.” (C)
- “I think it’s very good for this kind of project that you have a very wide range of requirements so it has narrowed down what we want to do.” (C)
- “It has given us some direction which I’m not sure whether we’d have had otherwise. Just because of the way we were working.” (D)
- “I think [TAPT] helped concrete [our] ideas.” (E)

Influencing the design process:

- “We always had TAPT in the back of our minds if it wasn’t written in stone […] it was part of the process.” (A)
- “A lot of the follow-on work was probably influenced by it.” (B)
- “Although we might not have specifically gone back to [TAPT] time and time again […] those ideas were always at the back of my mind when making these kind of decisions.” (D)

7.8.4 TAPT as an evaluative tool

The Blapr team originally intended to conduct a user-led evaluation of their system, but did not do this due to time constraints. However, Participant D did conduct a TAPT-based evaluation of Blapr. He remarked that it was an informal evaluation, whereby he Teased Blapr Apart based on his own knowledge and informal conversations with his teammates.

Regarding outcomes of the evaluation, he said he found that the abstract effects of Blapr were “really quite similar” to the abstract effects of signposts and tour guides, noting that he was not aware of these similarities until he conducted the TAPT process. As an example, he explained that QR codes (two-dimensional barcodes) are not superficially like signposts, but that they share properties such as being able to be placed “in strategic places around the site”. He said, “Although it’s not a signpost and it works in a completely different way to a signpost, it’s still got that aspect of it’s in this place, it tells you where you are […] it also directs you to places.” He concluded that Blapr’s properties “Tie in very well with the signpost and the tour guide”, saying “It’s actually quite surprising […] how close what we’ve come up with is […] to a signpost and a tour guide.”

Participant D noted that evaluation revealed “Interesting results whose value and relevance to the design of a product like Blapr is not immediately obvious.” He talked about the correlation between the abstract and literal effects of tour guides, signposts and Blapr, concluding that “These similarities are useful as a way to gauge the ease-of-use and quality with which the original brief has been satisfied.”

Participant D felt the evaluation was “Helpful in understanding [Blapr]” and that as an evaluative tool TAPT was “quite good”, saying “It helped me see that ‘yes actually, what we’ve done was quite close to what we originally intended’”.

Participant D identified various features of TAPT when used for an evaluation:

- Clarity: “It does help to clarify and order your thoughts.”
- Understanding: he felt that if the team hadn’t Teased Apart Google maps, tour guides and signposts as well as their own system “I wouldn’t have any idea how [Blapr] fit with any of them.

Participant D felt that TAPT sat naturally within the evaluation process as a “check” for ensuring the end product had delivered on its goals of providing users with a product to satisfy their needs in a familiar way, such that end users would be able to have experiences to match their expectations.

Participant D said he would use TAPT for evaluation again, provided he had used it at the initial design stage and therefore had the lists of effects for comparison: “With only the design stage, there is no concrete way to measure the effectiveness of designing a product around the gathered design data. Using just an evaluation stage, the only knowledge gained is the kind of experience that the product creates; there is nothing to say that the experience is the correct one.”

In summary, Participant D’s comments were very positive. This seems to indicate that using TAPT as part of a more sophisticated requirements testing process might be very effective: this is a potential area for future study.

7.8.5 Properties of TAPT

Case study participants identified other facets of TAPT, as outlined in Table 7-3.

Feature	Quote (participant, case study)
Efficient	"We actually spent very little time in a way […] you could do it again, because the number of hours [to apply it] is pretty low." (I, Location-based Tools)
Precise	"[TAPT is] very crisp [at translating experiences between domains]." (F, Spoken Web)
Clear	"[TAPT] will definitely help me design a better application […] I may not have addressed before [now] the experience of that social networking site. But I feel it has come up so clearly, it'll affect my approach […] There are some things I probably know but couldn't elucidate but it's come out in a very clear way." (F, Spoken Web)
Supports creativity	"It made us think outside the box." (B, Blapr) "It's also quite good for innovation so you could get ideas you wouldn't think of using this tool." "I think TAPT actually is quite good at leading you to get ideas like break them down into different sections and put down the different things that you think about." (C, Blapr)
Lends itself to repeated use	"It really would be useful to do this with many different groups actually, because the end result is very useful [and short] you could run it with a *lot* of groups and if you just looked at the end result you wouldn't have an insurmountable amount of data." (I, Location-based Tools)
Satisfying	Participant F (Spoken Web) would use TAPT again with user groups such as those at IBM ("If I had to do users with a similar profile to what we have done in Hursley I would use TAPT."). She said, "I am happy with the results" Participant F was positive about the way the study ran, saying the focus group worked "very well" and "better than I expected."
Structures / makes explicit an implicit process	"[TAPT] formalises […] makes explicit an implicit process." (B, Blapr) "You sort of do that anyway as you're designing but it's not a formal process, you look around and say 'this is doing this, this is doing that'. I think TAPT sort of formalises it and puts it into a structure so you … it does seem to help." (E, Blapr)
UX-oriented	"It was different […] The classic methods I've used in the past have been primarily about project management rather than actual 'let's design a system well'. They've been more to do with the technical aspects, how we're going to build it rather than what's the actual function going to be, are we heading in the right direction, is it going to do it in the right way, not just, is it going to do it? So it's a very good tool to gauge the appropriateness of the kind of thing that we're building. It's good." (D, Blapr)
User-generated terms more valuable than practitioner-generated terms	"The users themselves had to carry out the process [… in] a survey, I need to actually be asking the questions, prompting things, encouraging them to think, this process I like because you think for yourself. Probably [reduces the impact of] the bias of the person who's interviewing." (F, Spoken Web) "[Subjects] get the chance to explain what's going on […] people are essentially suggesting their own mechanisms." (G, Gaming) "This is [conducted] by the users […] perhaps I have prejudices that I'm not even aware of […] this method will maybe allow to cut straight through [those prejudices] […] it was very useful getting key words that users agreed upon and using them as a springboard to find the connections to do more analysis. I thought that was very interesting, a very good result there." (I, Location-based Tools)

Table 7-3 Features of TAPT

In summary, the practitioners identified various properties of TAPT:

- Allowing subjects to generate terms provided a rich vocabulary, and potentially reduced a bias which they, as practitioners, may introduce.
- TAPT was efficient, precise and clear.
- It supported creativity.
- It lent itself to repeated use (to corroborate data).
- It produced satisfying results.
- It made explicit and structured an implicit process.
- It supported the understanding of subtle aspects.
- It was oriented around user experiences.

7.8.6 Alternatives to TAPT

Participants were asked what alternative methods they might have used and how TAPT compared to these.

Few Blapr participants felt able to compare TAPT to other methods (with remarks such as “I haven’t used many design methods in the past”, “I'm not really aware of many design processes in general”, and “I’m not sure how much I’ve used really”). This is probably because they were graduate recruits to IBM and therefore early in their careers.

The Blapr team used TAPT in conjunction with Scenarios, and approved of this. Participants A and D felt the two methods were “quite complementary”, with Participant A talking about Scenarios’ user-focus and TAPT’s strength in “bringing in different ways of thinking about ways to design the product […] It’s different approaches. It gives you more than scenarios but I think they’re useful together.” Participant D talked about Scenarios as prompting the team to address user groups other than themselves, and the ability to apply that strength to facets of experience as identified by TAPT (“You need to consider not necessarily the stereotypical user of these things like a signpost and a tour guide […] a tour guide might have to give directions to someone who’s not necessarily in their group”).

Participant D said TAPT was “very different” to other approaches he had encountered, remarking that other design processes ask about what goals need to be achieved, saying they prompt questions such as: “have I done it?” He went on to say “TAPT is almost ‘how well have I done something’ it’s sort of a sliding scale not a discrete thing.”

Feature	Quote (participant, method to which TAPT is being compared)
Structures / makes explicit an implicit process	"Essentially I would have done TAPT in an informal process because to talk to people to understand what they want from a social networking application and then try and apply it to the [spoken web] … but it would be an unstructured way of doing things." (F, user survey) "My approach was similar to this, but not as structured. It was to try and work out what the key parts in each of the games were, as described in the literature, and then group them back together in some better genre types. Which is essentially what this should do." (H, literature analysis)
Clear	"[TAPT is a] very clear process." (F, user survey)
Experience oriented	"An interesting thing I found is the experience part [of TAPT]. Usually you focus a lot on functionality. You don't want to replicate exactly the functionality [here], but the experience." (F, user survey)
User articulation of concepts	"Especially since the users themselves had to carry out the process, this has helped introspection [when conducting a survey] I need to actually be asking the questions, prompting things, encouraging them to think […] I like [TAPT] because [subjects] think for [themselves]. It probably [reduces] the bias of the person who's interviewing." (F, user survey) "It's a bit more freeform in the sense that people are essentially suggesting their own mechanisms. I did mine starting with the mechanisms and seeing how people agreed, whereas these people are coming from nothing, generating their mechanisms themselves […] This let people bring out their own ideas on how things work." (G, online questionnaire)
Structured	"[Brainstorming is] not very specific… you put down whatever you think about and later on you look at it and say 'this might be a good idea' it's not like a system you can use." "[A literature survey] is not a structured approach […] TAPT adds structure which was not there before." She said this was "definitely" helpful and elaborated: "It's easier to distil what works in the web domain and what works in the voice domain when you have a formal and structured process as opposed to a random process because it's hard to quantify the claim when you're doing it on the basis of a literature survey." (F, literature survey)
Efficient	"You'd need to do a lot more work [with other methods] to come to similar conclusions." (F, literature survey) "I think getting the distilled key words and brief stuff is rather awesome actually. I think it's pretty efficient […] you'd quite likely get to the same [results as with textual analysis] but it'd be a far more round-about route." (I, textual analysis)

Table 7-4 Features of TAPT, compared to other methods

Participants identified contrasting features of TAPT as follows:

- Compared to brainstorming: structured (Participant C, Blapr).
- Compared to a user survey: makes explicit (and structures) an implicit process; clear; oriented around experiences; user articulation of concepts (Participant F, Spoken Web).
- Compared to a literature survey: structured, efficient (Participant F, Spoken Web).
- Compared to an online questionnaire: user articulation of concepts (Participant G, Gaming).
- Compared to a literature analysis: makes explicit (and structures) an implicit process (Participant H, Gaming).
- Compared to textual analysis: efficient (Participant I, Location-based Tools).

Table 7-4 shows features of TAPT which participants identified when comparing it to other methods.

7.8.7 Ways to use TAPT

Participant A reported the Blapr team chose to analyse signposts and tour guides because these items were related to location-based services. “We were going with things that we thought would be relevant, and different parts of what we might want to come up with -- things like a signpost and a tour guide”. This use of TAPT to analyse experiences like (but not identical to) the desired experience is interesting and was not predicted.

The Blapr team reported finding Teasing Apart easier with real objects (such as maps as opposed to the web-based Google maps) when trying to translate experiences to a web-based modality, because use of physical objects led to a stronger context switch. This battle to deconstruct virtual experiences mirrors comments from Participant G (Gaming), who felt that games could be viewed as “already distilled” (Section 7.4.4).

Participant B (Blapr) felt it was important to clarify the ways in which TAPT could be used, saying to him it had seemed that they had to take the outcomes of Teasing Apart and put them straight into Piecing Together to recreate the same experience. “At present its focus is on the Piece Together outcomes but the results from Tear Apart [sic] are equally if not more important and using those outcomes against a system that has begun its initial development phases would be useful." He felt there were other ways to use it:

- "See if you're on the right track with your current implementation."
- "As an evaluation tool."
- "To see if any of your other designs match some of these experiences from the first part."

Participant D said that if he were using TAPT again he would use it in partnership with a defined list of functional requirements for the system. He said, he would like to "Have that list of functional requirements and go through each after having done TAPT and say well how does this one apply to the system, how can we make sure that this requirement actually is not just done but we can say 'yeah it's done and it's good for these reasons because it's replicating these behaviours that people already expect'". He talked about the importance of usability of systems and how replicating known behaviours for users make it easier for those users, saying "Well it's what [they] expect." "You can get into the mindset of thinking what are the expectations of the user in terms of …they want this kind of information, how do we present that in a way that they already know and are familiar with?" He used an example of a digital camera saying "It has a whole range of amazing functions and it's brilliant and everything and it's all presented really well and I can understand how to change the aperture and things, but I've got no idea what changing the aperture does. Having functions that are known to me, those are the functions that I'd be much more likely to use and benefit from."

Participant D said he wasn't sure how well TAPT would work for end users to evaluate because it would be necessary to explain the process: "You'd have to describe the process and explain the reasoning behind each stage." He felt that using the TAPT framework but perhaps not the vocabulary to gather data from end users would work: "I think gathering user data without officially saying this is the framework we'll be using and this is what we're going to do, but having a conversation with them, asking opinions that would be useful to know for TAPT […] you do the evaluation with end users but are also there yourself to clarify that with TAPT … to think about these sort of things almost like as an interpreter to say 'this is what they've said and […] what they would've said if they were using TAPT language.'"

This mirrors a comment from Participant F (Spoken Web), who said that "[TAPT] works very well here, but in countries such as India and other developing countries we would have to come up with different processes." When asked to clarify, she said she felt TAPT was suitable for use in developing countries, but its presentation would need to be altered and simplified when working with, for example, illiterate end users.

7.9 Conclusions

This chapter presented four case studies of TAPT's use by practitioners, towards redesigning and understanding various experiences. Two studies involved using TAPT for design (and in one case, evaluation) and two involved using the method for analysis. Each case study had its own set of results:

1. Blapr: TAPT was used to successfully guide their design process, and also to evaluate the end product.
2. Spoken Web: TAPT was used to guide design processes which are still ongoing.
3. Gaming: TAPT was used to gain an initial set of analyses of genres of game, yielding early insights and paving the way for future studies.
4. Location-based tools: TAPT uncovered non-superficial similarities and differences between Gowalla and geocaching.

TAPT generally fulfilled participants' goals, which were oriented around understanding aspects of experience and building original and informed designs.

Participants used TAPT in various ways, applying it themselves, with focus groups and with anonymous subjects. They used it to understand similar experiences to the experience they were designing for; experiences they wished to redesign in new contexts; and experiences which they wished to investigate in their own right. Participants noted that their subjects did not always use TAPT as expected, and talked about using TAPT with many subjects to overcome this issue. Participants and their subjects took varying amounts of time to use TAPT, but appeared to speed up with experience and to work more quickly when given more support.

Participants felt that TAPT fitted in the design, evaluation and research processes, and commented positively about it as an aid to understanding, particularly of subtle and emotional aspects of experience. The Blapr case study found TAPT useful for supporting the replication of experience, and participants across all case studies identified various helpful facets of TAPT: they remarked upon it being efficient, precise, clear and satisfying. They commented on its usefulness for providing user-generated terms, supporting creativity, structuring an implicit process, and delving into user experience. They also remarked on its suitability for repeated use.

Participants compared TAPT to other methods, identifying contrasting features including its making explicit and structuring an implicit process, being experience-oriented, supporting user articulation of concepts, and being structured and efficient. Participants found TAPT and Scenarios to be complementary methods.

One participant used TAPT as an evaluative tool, and reported finding the process helpful for better understanding the product he evaluated.

Various approaches to using TAPT arose, including using it to analyse experiences *like* the experience of interest, finding TAPT easier with physical experiences, and applying TAPT to facets of an experience or to a list of project requirements.

These many and varied uses of TAPT are reminiscent of uses of the Unified Modelling Language (UML). UML is a flexible collection of methods (rather than a rigid process) used by Computer Scientists to describe the architecture and transactions of the systems they design and build. Often, Computer Scientists use the parts of UML which are best suited to the task at hand, rather than rigidly applying the whole language.

In a similar way, it would seem that TAPT can be used with a light touch, emphasizing the parts of the process most appropriate to the problem being tackled. TAPT allows a designer to focus on the user experience and choose how to incorporate those results into the rest of the engineering process. It can be used in an agile way.

The main future revision which could be made to TAPT based on these case studies would be to emphasise its flexibility, perhaps by including with TAPT material an outline of the ways in which it can be used, in terms of:

- Goal
    - Research-driven analysis: understanding experiences more deeply for research purposes, for example to evaluate the similarity of experiences.
    - Design-driven analysis: understanding experiences more deeply in order to redesign them more effectively, or in order to provide new products which encompass familiar experiences.
    - Evaluative analysis of a newly-built system or a work-in-progress, to compare experiences using that system to previously analysed experiences.
- Choice of experience to analyse
    - Modality: whether to choose physical or digital experiences, for example.
    - Choosing between directly relevant experiences and experiences *like* those of interest.
    - Considering facets of an experience: for example, working through a list of functional requirements of a system, or opting to consider only the thematic or the game-play elements of computer games.
- Logistics
    - Group or individual use.
    - Practitioner or end user application: if the latter, considering the way in which TAPT is presented – it does not have to be presented as a formal tool with analytical language.

The case studies provide encouraging evidence that TAPT is useful as a design, analysis and evaluative tool, which can be appropriated by a designer or analyst and used in a number of flexible ways.

Chapter 8 Conclusions

This chapter discusses the overall results presented in this thesis. First, there is a summary of the problem area and the research that was conducted, which includes reflections on how TAPT responds to current models of experience and a meta-reflection on the overall set of results. Next is a set of reflections on the process used: each stage of the evaluation is visited, and is followed by a meta-reflection on the evaluation process that refers to formal frameworks. The hypothesis and contributions are revisited before future work is discussed.

8.1 Reflections on Results

This thesis presented TAPT, a method for understanding user experiences and redesigning those experiences in new contexts. Although prior art included user-centred methods and approaches to working with novel, pervasive technologies, there was a dearth of methods to support the understanding and replication of user experiences. TAPT was developed in response to this need, and was motivated by a desire to broaden access to web-based social technologies by re-providing them via alternative modalities.

8.1.1 Initial application of TAPT (Chapter 4)

Results of the initial use of TAPT were that:

1. Applying TAPT in a systematic, repeatable manner resulted in useful analytical information and exposed emerging themes common across the items of social networking functionality.
2. Experienced effects were useful for distinguishing between superficially similar actions (such as public messaging and microblogging) that have different underlying effects.
3. Abstracting experiences with Teasing Apart made Piecing them Together easier. For example, it seems like a very big step to move from 'microblogging' to 'a scrolling display on a t-shirt'. By contrast, the steps from 'microblogging' to 'brief one-to-many communication' to 'a scrolling display' were much smaller and more logical.

It was clear at this stage that Teasing Apart is an inherently subjective process. One approach to improve its robustness might be to include multiple practitioners so that outputs are not the result of a single, potentially biased perspective.

TAPT's evaluation was performed in three phases:

1. A comparative evaluation, in which 43 Software Engineering professionals tackled design tasks using TAPT or one of two other methods.
2. An expert review, in which six domain experts evaluated the design artefacts produced in the comparative evaluation.
3. Four case studies, in which participants from industry and academia applied TAPT freely to their real-world work problems, and gave feedback on its usefulness.

These studies resulted in qualitative and quantitative feedback on TAPT's efficacy as a tool for analysis, design and evaluation, as well as insight into TAPT's features.

8.1.2 Comparative evaluation (Chapter 5)

This study provided strong evidence that TAPT is a structured method for thoroughly exploring facets of a given experience, including those emotional and social effects that might otherwise be overlooked. TAPT's primary weakness is that it is not used-focused, which can lead to less accessible designs.

Participants found the following aspects of TAPT to be helpful to a level with statistical significance:

- Overall usefulness (Table 5-15).
- Usefulness for understanding
    - emotional effects (Table 5-16).
    - perceptions and expectations (Table 5-20).
    - social effects (Table 5-18).
- Structure (Table 5-27).
- Workplace suitability (Table 5-30).
- Support of replicating perceptions and expectations (Table 5-25).

Other strengths identified were: supporting replication of experiences, fostering creativity and documentation of decisions. Other weaknesses were: its learning curve, formality, time required to apply it, and the necessity to choose which aspects to replicate.

Finally, participants rated TAPT as better than alternative methods at replicating superficial aspects, but not strongly enough to establish statistical significance: there is more discussion of this aspect in Chapter 5 (Section 5.4.2) and Chapter 6 (Section 6.3.2).

The comparative evaluation suggested that TAPT might be best used in tandem with a complementary method such as Scenarios, which is user-focused. The evaluation also suggested there might be benefits from applying TAPT to multiple points through the sequence of an experience, and to the experiences of multiple users within an experience.

Potential improvements that might be made to TAPT based on this study are:

- Providing lists of possible experienced effects, as a prompt.
- Encouraging reflection upon which experiential aspects designers may wish to omit.

8.1.3 Expert review (Chapter 6)

This review suggested that experts preferred artefacts produced with Unstructured Discussion (Table 6-16), although the evidence strongly suggests that this is due to a flaw in the normalisation process (Section 6.3.2), and possibly due to UD's support of superficial replication.

However, the expert review highlighted TAPT's strength as an analytical tool, particularly with respect to understanding abstract and emotional aspects of experiences. The experts also identified a disconnect between Teasing Apart analyses and Pieced Together designs, suggesting that more support needs to be provided in the Piecing Together process.

The expert review shed light on the impact of the method used in the comparative evaluation, providing insight on why picnic artefacts fell into two broad categories, why many wiki artefacts were wall-based, and showing that some analyses were 'led' by the examples given to participants.

Possible improvements to TAPT based on the feedback are:

- Take steps to combat the disconnect between Teasing Apart and Piecing Together, for example by encouraging practitioners to indicate in designs where they have included key effects from analyses.
- Encourage design teams to analyse the end domain as well as the starting domain.

8.1.4 Case studies (Chapter 7)

Four case studies were conducted:

1. Blapr: using TAPT to facilitate the design and evaluation of a multimodal messaging system for IBM.
2. Spoken Web: using TAPT to help design a voice-based system for IBM.
3. Gaming: using TAPT to gain an initial set of analyses of genres of game.
4. Location-based tools: using TAPT to understand non-superficial similarities and differences between two location-based tools, Gowalla and geocaching.

The case studies provided evidence that TAPT can be successfully used to support analysis, design and evaluation in real-world contexts, and that it is useful for supporting understanding and replication of experiences. Participants felt that TAPT was efficient, precise, clear and satisfying to use, and they said it was useful for providing user-generated terms, supporting creativity, structuring an implicit process, and delving into user experience. They also felt TAPT was suitable for repeated use.

When participants compared TAPT to other methods they spoke about its structure, ability to make explicit an implicit process, experience-orientation, support of user articulation of concepts, structure and efficiency.

Other insights were:

- TAPT can be applied in unexpected and sometimes inconsistent ways: using it repeatedly across multiple subjects may be a way to compensate for this.
- TAPT is quicker to use with prior experience of the method and support from an experienced TAPT user.
- TAPT and Scenarios are complementary methods.

The main conclusion from the case studies was that TAPT is suited to agile use: it can be used in ways that suit different activities (analysis, design, evaluation), experiences (modality, aspect of interest) and logistics (use by groups or individuals, practitioners or end users). TAPT was successfully used in the field in all of these ways.

8.1.5 TAPT and models of experience

Section 3.2.2 presented different frameworks of experience. TAPT was subject to a mixed methods evaluation, providing a deeper understanding of the method. This allows a discussion of how TAPT fits with those frameworks of experience.

Firstly, Forlizzi (Forlizzi, 2004) describes three types of user-product interactions (fluent, cognitive and expressive). TAPT could be used to analyse any of these areas: likely uses might be the analysis of 'fluent' usage towards replicating this experience elsewhere and of 'cognitive' usage (which is focused on understanding how to use a product) to try to understand how to improve existing products and make them more intuitive. 'Expressive' usage helps users form a relationship with a product, so TAPT's applicability here might centre on understanding the emotions (positive or negative) which result from these meaningful interactions.

Forlizzi also presents three types of experience context. These are 'experience' (a stream of 'self-talk' that occurs as we use products), 'an experience' (an event which can be articulated), and 'co-experience' (creating meaning and emotion through shared product use). TAPT could be applied to understand any of these arenas, although it seems more likely that practitioners might wish to analyse 'an experience' or 'co-experience' over 'experience'.

Hassenzahl (Hassenzahl, 2003) discusses two types of product attributes: pragmatic, utilitarian aspects, and hedonic attributes, centred on pleasure and emotional impact. As discussed through the course of this thesis, designer understanding of functional aspects of technology appears to be sound, and it is the hedonic aspects which are of strong interest in this arena: TAPT could be used to focus upon these.

Mahlke (Mahlke, 2005) identifies four dimensions of experience: perceived usefulness, ease of use, hedonic quality and visual attractiveness. TAPT could be used to explore all of these facets.

Wright and McCarthy (Wright, 2003) present a framework which compares four elements of experience, and six aspects about how people make sense of experiences. The four strands of experience are:

1. Compositional: what it is about, how it is composed, structure, consequences, explanations
2. Sensual: look and feel, sensations like thrill, fear, excitement, belonging, unease
3. Emotional: for example, anger, joy, disappointment, frustration.
4. Socio-temporal: the time and place of an experience

By breaking down experiences into their parts, TAPT investigates the structure of experiences in a specific way, allowing practitioners to gain insight into design and experiential aspects (the compositional strand). The abstract effects identified in TAPT allow investigation of sensual and emotional aspects (strands two and three). Finally, TAPT is applied to specific experiences which are situated in a time and place (strand four).

The second part of Wright and McCarthy's framework concerns how we make sense of experience. They discuss anticipating, connecting, interpreting, reflecting, appropriating and recounting experiences. This framework focuses on an experience as it happens over time, moving from anticipation beforehand through the experience and then on to post-experience reflections and recounting. Although TAPT focuses on an experience in a given time and place, the comparative evaluation in Chapter 5 included a question on how well the methods used supported understanding of people's perceptions and expectations: that is, their perceptions of the experience as a process. Participants rated TAPT as the best method in this area (and better than Unstructured Discussion with statistical significance), suggesting that although it looks at a specific experience, it encourages practitioners to consider expectations beforehand and reflections afterwards. Further exploration of this area would yield greater insight.

It can be seen that TAPT can be widely applied to many facets of experience, according to different models. This is a further reflection of its suitability for agile use, as revealed in Chapter 7.

8.1.6 Further reflections

A theme through all three studies was that participants tended to struggle somewhat with the distinction between literal and abstract experienced effects. The intent in separating effects into two parts was to scaffold the process by making the steps smaller and less difficult and encourage practitioners to consider tangible, concrete effects (a generally simpler task) first, giving a gentle start to the process of analysing experienced effects.

It seems likely that this aspect could have been improved with clearer examples: for example, 'sharing past experiences' was presented to participants of the comparative evaluation as an example literal effect of photo-sharing on Facebook, when it could be argued that this is no more concrete than 'openness about past experiences' or 'reminiscence', which were presented as abstract effects of the same experience.

However, the problem appears to be greater than that. Potential alternatives to 'literal' and 'abstract' might be 'experiences' and 'consequences'; 'what you do' and 'what you think or feel'; 'physical' and 'abstract or virtual' – but none of these cover the divide clearly. The intent was to capture the difference between measurable outcomes which change in the real world (physical changes such as food being eaten or money spent, along with virtual changes such as text appearing on a webpage or a social link being made between two online entities) and intangible changes involving feelings. 'External changes' and 'internal changes' might differentiate these types of outcome more accurately.

Participant feedback suggests that although identifying abstract effects was often difficult, it was also worthwhile.

8.2 Reflections on Process

The overall sequence of the TAPT evaluation was:

1. Initial application to facets of social networking (Chapter 4)
2. Comparative evaluation (Chapter 5)
3. Expert review (Chapter 6)
4. Case studies (Chapter 7)

The above referenced chapters provide reflections on the individual methods used. This section summarises those insights and provides holistic thoughts on the overall process taken to evaluate TAPT.

8.2.1 Initial application of TAPT (Chapter 4)

Applying TAPT to facets of social networking proved that it is possible to yield useful analytical results with the method. However, this initial usage only showed that TAPT could be used by its creator and not necessarily anyone else. This was a prompt for broader evaluation, although TAPT's focus on subjective experiences meant the evaluation had to be approached with care.

8.2.2 Comparative evaluation (Chapter 5)

The comparative evaluation demonstrated that TAPT could be used systematically by Software Engineering professionals and resulted in various insights into TAPT's efficacy. There were two potential weaknesses with the experimental method: firstly, that participants' assessments of their own designs were likely to be biased, and secondly that the experiment as a whole was run in an artificial laboratory context with TAPT applied to a fictional scenario. The subsequent expert review was intended to address the first weakness by gaining more objective insight into the resultant artefacts, while the case studies examined TAPT's use in the field.

The results of the comparative evaluation provide an example of how people can be constrained by questionnaire formats to give inaccurate responses: participants appeared to dislike rating TAPT as 'bad' at replication of surface elements when they had chosen not to replicate such elements. They felt that saying TAPT was 'very bad' implied that they wanted to achieve this goal but couldn't, rather than choosing not to pursue that goal. (This was indicated by comments such as "We deliberately tried to be different from the wiki"; "We intentionally changed them all", and participants asking whether it was 'ok' to tick 'very badly' when they were happy with their design: see Section 5.3.3).

Additionally, as noted in Section 5.4.6, participants gave different comparative ratings of TAPT's efficacy at replicating experiences in different questions. This difference could be due to a number of factors, either individually or in combination:

- Changing perceptions of participants over the course of the study.
- Changing interpretations of what is meant by 'replication' over the course of the study or depending on how the question is asked.

8.2.3 Expert review (Chapter 6)

The expert review resulted in strong insights into TAPT and into the method used in the comparative evaluation. Findings suggested that some participants were led by the examples they were given, suggesting that despite the pilot studies conducted to hone the experimental materials, they were still not quite right. Evidence from the expert review suggested that the comparative evaluation would have yielded fairer results had participants using unfamiliar methods been allowed a trial run of the new method first, to familiarise themselves with it.

Interviewing pairs of experts in order to ascertain their level of agreement worked well. Reflecting on the overall method yielded insights into how the method could have been honed further, in two respects:

A flaw with the experiment was the impact that the normalisation process had on results. Design artefacts were normalised to mask the method by which they were produced: the approach largely consisted of removing the analytical material from structured methods like TAPT and Scenarios. Removing this material appeared to do those artefacts a disservice, evidenced by the low ratings of doctored artefacts and expert comments on their simplicity, in contrast to expert praise of analysis.

The second problem with the experiment was that one of the six experts was from a substantially different background to the rest (Expert C was a Sociologist, in contrast to the other experts who were from engineering backgrounds). Every expert had a strong understanding of the domain of the task they were asked about, but Expert C's discomfort with technology and engineering appeared to leave her distanced from the design artefacts which concerned a web-based recreation of picnicking: she found them difficult to evaluate. The lesson from this is that expert knowledge must encompass the technological domain as well as the application domain.

One further problem with the expert review was the awkwardness of the question intended to elicit opinions on whether blind artefacts replicated experiences. The question was:

1. Does the artefact translate the experience? And is that translation of the deeper, underlying experience – for example, of emotional or social aspects of the experience – or is it a translation of more superficial, design elements?

With hindsight, asking experts two separate questions would have yielded clearer responses:

1. Did this artefact replicate the superficial, design elements of the experience of [using a wiki | microblogging | picnicking]?
2. Did this artefact replicate the deeper, emotional or social aspects of [using a wiki | microblogging | picnicking]?

8.2.4 Case studies (Chapter 7)

The case studies yielded strong results. The advantage of using case studies was their ability to yield results grounded in industrial and academic practice: TAPT was applied to real world problems, rather than artificial scenarios in the laboratory. The case studies were able to take advantage of knowledge gained from earlier experiments: for example, the Blapr team successfully used TAPT and Scenarios in a complementary fashion, a possibility first raised during the comparative evaluation.

The main disadvantage of the case studies arose from the inevitable lack of control involved in equipping participants to freely use the method as they pleased. The Spoken Web and Gaming case studies were still ongoing at the time of writing, and although both yielded useful results, more data would have been available had interviews been conducted later. This problem is inherent to the case study approach, and little can be done about it beyond placing interviews at appropriate project milestones.

An additional disadvantage of case studies was that the researcher was not necessarily present when TAPT was used: for example, the information on how long Blapr participants spent applying TAPT came from the participants' recollections two weeks after they first applied the method. This could have been somewhat mitigated: for example, at the beginning of the Blapr study, participants could have been asked to keep more detailed notes of how long it took them to apply TAPT. The lesson here is that researchers should be aware of exactly what material they want from case studies before the studies begin, so that participants can be asked to be aware of any particular facets of interest when they conduct work away from the research laboratory.

8.2.5 TAPT's evaluation and formal evaluation frameworks

Section 3.3 discusses the evaluation of design methods. It included a discussion of rigour and relevance, and case studies. It also presented frameworks from Kitchenham (Kitchenham, 1996) and Furniss. This section discusses how TAPT's three-pronged evaluation fits with the above aspects.

Section 3.3 described Fallman's (Fallman, 2010a) discussion of rigour and relevance. 'Rigour' emerges from research's validity and reliability: the results in this thesis appear to conform to these measures, as they explain the TAPT process and within the studies conducted, appeared to be repeatable. For example, case study results correlated with one another.

In terms of 'relevance', Fallman focuses on Keen's criteria (Keen, 1991). These are that research addresses problems important to professionals, can be used by practitioners, addresses timely issues, and is understandable by practitioners. The successful execution of the four case studies presented in this work suggests that TAPT is indeed relevant to current problems and can be used by professionals to address those problems.

Section 3.3 included a discussion of the use of case studies to investigate phenomena within real-life contexts and answer "how" and "why" questions. This thesis presents the use of case studies to investigate how TAPT would be used by practitioners, and which features of the method they would identify as particularly helpful or unhelpful. Shneiderman (Shneiderman, 2007) discusses the use of multiple case studies to replicate findings with diverse users and problems: this thesis presents the use of four studies, the results of which do correlate. Further case studies would enable generalisation of the results with greater confidence.

Kitchenham (Kitchenham, 1996) presents nine approaches to evaluation of Software Engineering methods and tools. Table 8-1 shows how the evaluation of TAPT fits with this group of approaches.

Method	**Description**	**Whether this method was used**
Quantitative experiment	Investigate the quantitative impact of methods via a formal experiment	Yes (comparative evaluation, Chapter 5)
Quantitative case study	Investigate quantitative impact of methods via a case study	No (the case studies did not yield quantitative data)
Quantitative survey	Investigate quantitative impact of methods via a survey	No (surveys are only applicable when methods have been in use for some time)
Qualitative screening	A feature-based evaluation conducted by an individual who determines the features to be assessed and their rating scale	Yes (the literature review in Chapter 3 revealed that existing methods do not support replication of experiences, and rarely support understanding emotional and social aspects)
Qualitative experiment	Feature-based evaluation carried out by a group of potential users who try out the methods on typical tasks before making their evaluations	Yes (comparative evaluation, Chapter 5)
Qualitative case study	Feature-based evaluation performed by someone who has used the method on a real project	Yes (case studies, Chapter 7)
Qualitative survey	A feature-based evaluation by people with experience of the method	No (surveys are only applicable when methods have been in use for some time)
Qualitative effects analysis	A subjective assessment of the quantitative effect of methods based on expert opinion	Yes (expert review, Chapter 6)
Bench-marking	Running standard tests using alternative tools / methods and assessing the relative performance of the tools	No (not applicable: benchmarking is relevant for tools but not methods such as TAPT)

Table 8-1 TAPT's evaluation according to Kitchenham's set of methods

As can be seen, the mixed methods evaluation of TAPT included five of the nine suggested methods, excluding: the two types of survey, which are for use when methods have been in use in organisations for some time; benchmarking, which is intended to measure technological tools rather than methods for which human expertise is required; and quantitative case studies, because the case studies conducted were intended to garner qualitative insights rather than more numeric data.

As such, TAPT's evaluation was, in terms of Kitchenham's set of possible approaches, comprehensive.

As described in Section 3.3, Furniss (Furniss, 2008) suggests four important contexts of usability practice that influence the use of methods. These are:

1. Fit with current working practice.
2. Fostering relationships with clients and colleagues.
3. Supporting communication of ideas.
4. Other aspects of working practice, such as reporting and visibility within communities.

It is of note that the final facet, “other aspects of working practice”, relates to understanding how a method fits with context of real usage, and that it can take five to ten years to gain a full understanding of this aspect. Cognitive Walkthroughs, for example, evolved a great deal between their original publication (Lewis, 1990) and the publication of the procedure which was considered definitive (Wharton, 1994); a further six years passed before context of working practice prompted publication of a ‘streamlined’ version which better supported these four contexts (Spencer, 2000).

The following subsections discuss how results from the experiments described in this thesis discuss TAPT in terms of a) fit with current working practice and b) fostering relationships and supporting communication.

8.2.5.1 Fit with current working practice

One facet of working practice is the suitability of a method’s output for workplace use. The comparative evaluation (Chapter 5) asked participants about this (Section 5.4.5): TAPT was rated at 3.5/5, Scenarios at 3.6/5 and Unstructured Discussion at 2.7/5. TAPT’s rating was statistically significantly better than the rating of UD, and participants commented positively about its intuitive layout and the ability to trace decisions, with some participants suggesting the addition of space for flow charts and diagrams.

The case studies presented in Chapter 7 also addressed this area. All participants were asked whether they felt TAPT sat naturally within the processes they applied (whether design, analytical or evaluative): Section 7.8.1 reports on their comments, but in summary seven of the nine participants felt TAPT sat naturally within their work processes. The two exceptions were both on the Blapr team: one felt that their project did not involve replicating experience (making TAPT less relevant), while the other felt that the team had ‘failed’ to define a problem, resulting in a less good fit.

The case study responses do not constitute statistically significant data, but in conjunction with evidence from the comparative evaluation they suggest that TAPT fits with current working practice when applied to appropriate problems.

8.2.5.2 Fostering relationships and supporting communication

The comparative evaluation presented in Chapter 5 included two questions relevant to this area, asking participants how well their methods scaffolded discussions about the problem, and how suited methods' output was to workplace use. As noted in the previous section, TAPT output was rated highly in the context of workplace use, and its ratings were significantly better than those of Unstructured Discussion.

Participants also responded positively to the question on how well TAPT scaffolded discussion, rating it at 4.3/5 (compared to 4/5 for Scenarios and 2.9/5 for UD); again, TAPT's ratings were better than those of UD with statistical significance. Participants commented on the usefulness of TAPT's structure, although some brought up the issue of a disconnect between the Teasing Apart and the Piecing Together phases; these responses correlated with observations from the expert review about a disconnect between the two phases (see Section 6.4.3).

As well as confirming observations about difficulties using Teased Apart analyses to Piece Together designs, results from the expert review (Chapter 6) strongly suggested that *full* TAPT artefacts succeed in communicating practitioner decisions and thought processes (Section 6.4.2), but truncated artefacts do not (Section 6.3.2).

The case studies in Chapter 7 provide further evidence for TAPT's ability to foster communications and relationships. As described in Section 7.7.1, TAPT was frequently used collaboratively: in three of the four studies (Blapr, Spoken Web and Location-based Tools), TAPT was only by groups. TAPT was used around a whiteboard and via focus groups.

Section 7.8.5 described features of TAPT identified by participants: these included it being 'precise' and 'clear', both facets to foster communication. Multiple participants commented on the value of user-generated terms over practitioner-generated terms: this demonstrates TAPT facilitating communication between users and practitioners.

Section 7.8.6 described features of TAPT which participants identified when comparing the method to alternative approaches which they might use in the course of their work. Participants said that TAPT was ‘clear’ and again returned to the value of end users articulating concepts.

Finally, Section 7.8.7 discussed ways to use TAPT: here, Participant D (Blapr) and Participant F (Spoken Web) expressed opinions that the TAPT framework could be used with broad sets of end users, but that its presentation might need to be tailored to more diverse audiences, for example if working with end users who may be illiterate. This thesis has shown that TAPT as it currently exists is suited to use by Software Engineering professionals, but these comments suggest the method could be adapted to broader audiences.

In summary, the evidence presented in the comparative evaluation, expert review and case studies suggests that TAPT fosters relationships and supports communication.

8.2.6 Further reflections

As observed, existing literature draws attention to a perceived ‘pressure’ in creativity and design research to seek quantitative results, and a tendency to assume that these correlate with rigour (Shneiderman, 2007) (Fallman, 2010a). Meanwhile, Yin (Yin, 2008) notes a concern that case studies provide little basis for scientific generalisation.

This work responded to these issues through the use of mixed methods, which also enable multiple perspectives and ‘triangulation’ of data (Shneiderman, 2006) (Isomursu, 2007). Quantitative results with statistical relevance were gained via the comparative evaluation and (to an extent) the expert review. All three experiments yielded qualitative results.

The use of three mixed methods experiments enabled the acquisition of results that were both broad (exemplified by the 17 TAPT artefacts generated by 35 individuals in the comparative evaluation) and deep (the four case studies of TAPT’s use). The use of case studies yielded results which were grounded in industrial and academic practice, guided by the results and lessons learned from the prior studies: for example, the Blapr team used TAPT in conjunction with Scenarios based on results from the comparative evaluation.

Laboratory-based studies yielded information about the *way* in which TAPT supported understanding and replication (through a large set of answers to detailed questions on facets of these two aspects) while the case studies enabled examination of *how* TAPT could be used in the field (providing evidence of TAPT's usefulness in a series of different contexts).

Although the expert review examined, and resulted in, further insight into the designs from the comparative evaluation, it was not possible to conduct a similar process with the questionnaire responses from the comparative evaluation. Although participant responses (particularly those with statistical significance) were clearly useful, the sometimes mixed responses on how well TAPT did at replicating experience provide something of a warning about taking people's ratings of task success at face value: for example, as described in Section 5.4.2, participants clearly interpreted the meaning of 'replication' in different ways.

Perhaps the main limitation of the results presented in this thesis is the difficulty in gaining insight into the levels of superficial and deep replication of experiences. As discussed, there were problems in eliciting feedback on this aspect in both the comparative evaluation and the expert review. Additionally, the four case studies did not yield a great deal of data on how TAPT can be used to replicate experiences. Two of the four studies (Gaming, Location-based Tools) were analytical and did not involve the replication of experience, while one of the other two (Spoken Web) is still unfinished. The Blapr study was therefore the only case study to implement a system intended to replicate an experience. Feedback on this was highly positive, but it would have been helpful to corroborate the results with at least one other study.

It is of note that questions about replication of experience on different levels are inherently complex. It would appear that superficial disparities between original and reconstructed experience can lead people to overlook the more fundamental similarities that have been achieved by a comprehensive process of redesign, and this is an issue which would remain regardless of the finer aspects of methodology.

Overall, the use of multiple evaluation methods offered the benefits described as 'triangulation' when discussing the use of mixed methods: just as using a combination of quantitative and qualitative measurements enables researchers to gather richer insights, so using a combination of laboratory-based studies and fieldwork opens up further types of research finding.

8.3 Thesis and Contributions

This thesis identified an apparent lack of support from current tools and techniques for understanding and replicating experiences, particularly in terms of emotional and social effects. It described the motivation, development and evaluation of TAPT, a method built in response to this gap. TAPT was shown to be a structured method for thoroughly exploring facets of a given experience, including emotional and social effects that might otherwise be overlooked. Identifying these effects is essential for understanding an experience in an abstract sense, as an intermediate step for redesigning the experience for a different context.

The strengths and weaknesses of TAPT have been identified and discussed in detail, as have potential improvements to it. These results were shown through the use of a large-scale comparative evaluation, an expert review of the output of that evaluation, and case studies showing TAPT's use in practice. Results were both qualitative and quantitative in nature, enabling triangulation on important and unexpected aspects.

8.3.1 Thesis

This thesis addressed the following hypothesis:

> *A systematic process of user experience deconstruction (that identifies superficial, abstract and emotional elements) can improve the critical analysis of user experiences, and in a software development context better scaffold the initial design process or evaluation of user experiences compared to existing methods.*

The hypothesis was broken down into five research questions:

1. Is the TAPT process of user experience deconstruction systematic?

Results strongly suggest that TAPT is systematic. Chapter 5 and Chapter 7 describe TAPT's use by 43 participants in a laboratory-based study and nine participants in case studies, showing it lends itself to repeated use across various contexts.

Additionally, the laboratory study resulted in statistically significant findings on TAPT's systematic nature: participants found TAPT more suitable than Unstructured

Discussion for workplace use (Table 5-30) and for scaffolding structured conversation (Table 5-27).

2. Can a TAPT analysis identify superficial, abstract and emotional elements?

Results strongly suggest that TAPT helps users identify abstract and emotional elements. Results also suggest that TAPT is helpful for identifying superficial elements.

Chapter 5, Chapter 6 and Chapter 7 provide qualitative comments on the identification of these elements and statistically significant results on TAPT's overall usefulness (Table 5-15) and its usefulness for understanding emotional effects (Table 5-16), perceptions and expectations (Table 5-20) and social effects (Table 5-18). Study participants rated TAPT as better than alternative methods at replicating superficial aspects, but without statistical significance: there is more discussion of this aspect in Chapter 5 and Chapter 6.

3. Can a TAPT analysis be used to scaffold the initial software design process?

Chapter 5 presents a comparative evaluation of TAPT's use towards early design thinking and Chapter 6 gives an evaluation of the design artefacts developed in the first experiment. Chapter 7 includes two case studies of TAPT's use for design problems (Blapr, Spoken Web).

Statistically significant results from the comparative evaluation were that participants found TAPT better than Unstructured Discussion at replicating expectations and perceptions (Table 5-25). The expert review suggested that experts preferred artefacts produced with Unstructured Discussion (Table 6-16), although the evidence strongly suggests that this is due to a flaw in the experiment method (Section 6.3.2).

The evidence strongly suggests that TAPT provides very good support for conducting analysis (see question 2) and some support for conducting design. As such, it scaffolds the early design stages of the software process, particularly in its analytical phase.

4. Can a TAPT analysis be used to evaluate user experiences to support software development?

Chapter 7 includes a case study (Blapr) where TAPT is successfully used to compare user experiences towards evaluating a software system. Although this is only one case study, it provides evidence that TAPT can be used in this way.

5. Can a TAPT analysis be used for critical analysis of user experiences?

Two case studies (investigating gaming and location-based tools) in Chapter 7 demonstrate the application of TAPT for critical analysis of user experiences in a research context, showing that it is possible to use TAPT in this sense. One case study (gaming) acquired an insufficient number of TAPT analyses to conduct a meta analysis, but participants felt the analyses would be of use if there were more of them. The second case study (location-based tools) was successful, with the participant strongly stating that TAPT was a useful tool in this context.

8.3.2 Contributions

This thesis documented three key contributions made to the field of user experience:

1. The TAPT method, described in Chapter 4

TAPT addresses the gap in the field of software design which is identified in Chapter 1, exemplified in Chapter 2 and explained in Chapter 3: it enables the understanding of user *experiences* and scaffolds the redesign of these for new contexts.

2. A three-part mixed methods analysis of TAPT, presented in Chapter 5, Chapter 6 and Chapter 7

This evaluation of TAPT provides quantitative and qualitative evidence of TAPT's efficacy in terms of facilitating understanding and replication of experiences. It also provides participant feedback about TAPT's qualities (for example, its structured nature) and ways in which it can be applied.

3. Reflections on approaches to analysing Software Engineering design processes such as TAPT. These are given within each evaluative chapter. This final chapter presented an overall, holistic reflection.

Design is a creative, qualitative process, and understanding user experiences is a subjective and difficult task. As such, evaluating a process intended to support design through understanding experiences is non-trivial. The reflections and lessons learned from evaluating TAPT can be more broadly applied by practitioners in the field.

8.4 Future work

The three evaluative studies each suggested potential improvements to TAPT. Specific possible changes identified by this work are:

- Providing lists of possible experienced effects, as a prompt.
- Encouraging reflection upon which experiential aspects designers may wish to omit.
- Allowing practitioners to include diagrams and sketches in the Piecing Together phase.
- Take steps to reduce the disconnect between Teasing Apart and Piecing Together, for example by encouraging designers to indicate in designs where they have included key effects from analyses, or what effects they have not included and why.
- Encouraging design teams to analyse the end domain as well as the starting domain.

The first target for revision of the TAPT process based on the research conducted so far would be to improve the Piecing Together phase. This is important because although the evidence suggests that Piecing Together does support the redesign of experiences, it does not appear to enable practitioners to fully use the powerful analyses that they can generate with Teasing Apart.

Other future investigations might involve:

- Examining how practitioners apply the Teasing Apart phase when they do not know the end domain into which the experience will be Pieced Together.
- Investigating how groups apply TAPT, and ideal group size and composition.
- Exploring how to apply TAPT a) repeatedly (at points through one experience and for multiple users in one experience) and b) in conjunction with other methods such as Scenarios.
- A comparative evaluation of TAPT’s usage in the different ways exemplified in the case studies.

Another important task for future research would be to apply TAPT towards solving the problem described in the introduction of this thesis, that of enabling broader access to social technologies which are currently web-based. The evidence in this thesis strongly suggests that by applying TAPT to analyse and understand existing experiences of social networking technologies, researchers and designers will be well-placed to redesign them for use in novel contexts.

TAPT has been evaluated in the context of UK Software Engineering (excepting the two case studies of its use in British and Norwegian universities): at the moment, the results presented in this thesis can only be interpreted in this context. Further evaluations of TAPT's use in broader contexts would offer greater insights into its applicability. For example, archaeologists conduct a process called phenomenology, which involves using sensory input to interpret a site, attempting to separate their 21st century perspective from what they see. Current software helps them build virtual models, but they report that this software is unhelpful: TAPT might be used with domain experts such as archaeologists to design better software.

8.5 Conclusion

This thesis was motivated by a gap in existing tools and techniques with respect to understanding and building designs oriented around user experience. TAPT is a method developed in response to that gap: the evidence demonstrates that it is a powerful tool for understanding experience and that it supports the design process. Although scaffolding the creative leap is no easy task, TAPT provides stepping stones to support the process of redesign, and makes that process recordable and replicable.

TAPT has been successfully used in industry and academia for design, research and evaluation. Evidence suggests it can be used in an agile fashion which encompasses, but is not limited to, Software Engineering. Meanwhile, Tim Berners-Lee has suggested that semantic web technologies could provide a mechanism for implementing the redesigned visions built using TAPT (Berners-Lee, 2010).

Industrial and academic indicators suggest that this work is timely. For example, an employee of IBM Canada engaged in user research remarked on the importance of helping developers be empathic towards their end users: "Offering design frameworks which help [developers] become empathetic [sic] and aware of the target market will only serve to help the design playing field" (Neable, 2009). She felt this was of particular importance with IBM moving into areas such as Business Process Management, where end users were from very different backgrounds and skill sets to IBM software developers.

As technological social systems continue to rise in prominence, it is necessary to support the design of strong and innovative social systems in broader contexts, a task to which TAPT is highly suited. Meanwhile, the rise of disciplines such as Web Science demonstrate the need to understand the world around us in increasingly holistic ways: TAPT represents a way to achieve that understanding, and to support the design of the next generation of digital tools.

References

Abram, C., Pearlman, L. 2008. The Many Faces of Facebook. *Facebook For Dummies.* Wiley.

Anderson, K., Jacobs, M., Polazzi, L. 2003. Playing Games in the Emotional Space. *In:* Blythe, M., Overbeeke, K., Monk, A.F., Wright, P.C. (eds.) *Funology.* Springer.

Andrews, C. 2001. Unified Communications Systems. *The ACM Student Magazine.* Fall 2001 edition.

Ark, W. S., Selker, T. 1999. A look at human interaction with pervasive computers. *IBM Systems Journal,* 38**,** 504-507.

Atwood, M. E., Mccain, K.W., Williams, J. C. 2002. How does the Design Community think about Design? *Designing Interactive Systems.* London, UK: ACM Press.

Bardzell, J., Bardzell, S. 2008. Interaction Criticism: A Proposal and Framework for a New Discipline of HCI. *CHI.* Florence, Italy: ACM.

Berners-Lee, T. 2010. *RE: Conversation about applications of TAPT. (Personal communication)*

Bernstein, M. 1998. Patterns of Hypertext. *Hypertext.* Pittsburgh, USA: ACM.

Blythe, M., Wright, P., McCarthy, J., Bertelsen, O. Theory and Method for Experience Centered Design. CHI, 2006 Montreal, Quebec, Canada. ACM, 1691 - 1694.

Boehner, K., Depaula, R., Dourish, P., Sengers, P. 2005. Affect: From Information to Interaction. *Critical computing: between sense and sensibility.* Aarhus, Denmark.

boyd, D. M., & Ellison, N. B. 2007. Social Network Sites: Definition, History, and Scholarship. *Journal of Computer-Mediated Communication,* 13.

Branco, P. 2001. Challenges for Multimodal Interfaces Towards Anyone Anywhere Accessibility: A Position Paper. *In Workshop on Universal Accessibility of Ubiquitous Computing: Providing for the Elderly.*

Brand, S. 1994. *How Buildings Learn,* New York, Viking.

Brandeis, L., Warren, S. D. 1890. The Right to Privacy. *The Harvard Law Review,* 4**,** 193–220.

Brown, M., Weiser, R., Gold, J.S. 1999. The origins of ubiquitous computing research at PARC in the late 1980s. *IBM Systems Journal,* 38.

Buchenau, M., Suri, J.F. 2000. Experience Prototyping. *DIS.* New York, USA: ACM.

Callon, M. 1986. Some Elements of a Sociology of Translation: Domestication of the Scallops and the Fishermen of St Brieuc Bay. *In:* Law, J. (ed.) *Power, Action and Belief: A New Sociology of Knowledge* Routledge.

Candy, L., Edmonds, E. 1999. Introducing Creativity to Cognition. *Creativity and Cognition.* Loughborough, UK: ACM.

Cheverst, K., Clarke, K., Dewsbury, G., Hemmings, T., Kember, S., Rodden, T., Rouncefield, M. 2003. Designing assistive technologies for medication regimes in care settings. *Universal Access in the Information Society,* 2**,** 235-242.

Cheverst, K., Cobb, S., Hemmings, T., Kember, S., Friday, A., Phillips, P., Procter, R., Rodden T., Rouncefield, M. Design With Care. EDRA32 2001, Workshop Symposium on "Adding the user into the design of the home in the Digital Age", July 2001 Edinburgh.

Cody, M., Dunn, D., Hoppin, S., Wendy, P. 1999. Silver surfers: Training and evaluating internet use among older adult learners. *Communication Education,* 48**,** 269-286.

Colvin, J., Chenoweth, L., Bold, M., Harding, C. 2004. Caregivers of Older Adults: Advantages and Disadvantages of Internet-based Social Support. *Family Relations,* 53**,** 49-57.

Cooper, A., Reimann, R., Cronin, D. 2007. The Foundations of Design: Scenarios and Requirements. *About Face 3: The Essentials of Interaction Design.* Wiley.

Curtis, G., Vertelney, L. 1990. Storyboards and sketch prototypes for rapid interface visualization. CHI Tutorial: ACM Press.

Daniel, K., Cason, C., Ferrell, S. 2009. Assistive Technologies for Use in the Home to Prolong Independence. *PETRA.* Corfu, Greece: ACM.

Davis, F. D. 1989. Perceived Usefulness, Perceived Ease of Use, and User Acceptance of Information Technology. *MIS Quarterly,* 13**,** 319-340.

de Roure, D., Hey, T., Trefethen, A. E. 2005. *Where the Grid meets the Physical World - Research Issues in Grid and Pervasive Computing* [Online]. Available: http://www.semanticgrid.org/documents/gridperv3.pdf.

Desmet, P. 2003. Measuring Emotion: Development and Application of an Instrument to Measure Emotional Responses to Projects. *In:* Blythe, M., Overbeeke, K., Monk, A. F., Wright, P. C. (eds.) *Funology.* Springer.

Dewsbury, G. 2001. The social and psychological aspects of smart home technology within the care sector. *New technology in human services,* 14**,** 9-17.

Dix, A. 2003. Deconstructing Experience - pulling crackers apart. *In:* M. Blythe, K. O., A. F. Monk, P. C. Wright (eds.) *Funology: From Usability to Enjoyment.* Dordrecht, the Netherlands: Kluwer.

Donaldson, J., Evnin, J., Saxena, S. 2005. ECHOES: Encouraging Companionship, Home Organization, and Entertainment in Seniors. *CHI.* Portland, Oregon, USA.

Dowdall, A., Perry, M. The Millennium Home: Domestic Technology to Support Independent-living Older People. The 1st Equator IRC Workshop on Ubiquitous Computing in Domestic Environments, 2001 Nottingham.

Edwards, W. K., Grinter, R. At Home with Ubiquitous Computing: Seven Challenges. Ubicomp2001, 2001 Atlanta, GA, USA.

Fallman, D., Stolterman, E. 2010a. Establishing Criteria of Rigor and Relevance in Interaction Design Research. *create10.* Edinburgh, UK.

Fallman, D., Waterworth, J. 2010b. Capturing User Experiences of Mobile Information Technology with the Repertory Grid Technique. *Human Technology,* 6**,** 250-268.
Forlizzi, J., Battarbee, K. 2004. Understanding Experience in Interactive Systems. *DIS2004.*
Foth, M. 2006. Analyzing the Factors Influencing the Successful Design and Uptake of Interactive Systems to Support Social Networks in Urban Neighborhoods. *International Journal of Technology and Human Interaction,* 2**,** 65-79.
Frazer, A., Argles, D., Wills, G. 2008. Demystifying the Educational Benefits of Different Gaming Genres. *ALT-C 2008: Rethinking the digital divide.* Leeds, UK.
Fried, C. 1968. Privacy. *Yale Law Journal,* 77**,** 475–493.
Furniss, D., Blandford, A., Curzon, P. 2008. Towards Maturing Usability Practice in Website Design: Grounding how practitioners work to inform research requirements. *In:* E. Law, E. H., G. Cockton & J. Vanderdonckt (eds.) *Maturing Usability: Quality in Software, Interaction and Quality.* London: Springer.
Gamma, E., Helm, R., Johnson, R., Vlissides, J. 1995. *Design patterns : elements of reusable object-oriented software*, Addison Wesley.
Gaver, B., Dunne, T. Pascenti, E. 1999. Design: Cultural probes. *Interactions,* 6**,** 21-29.
Gaver, W., Boucher, A., Martin, H. 2003. Evaluating the Double-Deck Desk.
Gaver, W., Bowers, J., Boucher , A., Gellerson, H., Pennington, S., Schmidt, A., Steed, A., Villars, N., Walker, B. 2004. The Drift Table: Designing for Ludic Engagement. *CHI.* Vienna, Austria.
Gaver, W., Bowers, J., Boucher, A., Law, A., Pennington, S., Villar, N. The history tablecloth: illuminating domestic activity. Designing Interactive systems, 2006 University Park, PA, USA.
Gero, J. S. 2002. Computational Models of Creative Designing Based on Situated Cognition. *Creativity and Cognition.* Loughborough, UK: ACM.
Gibbs, G. 2004. Searching for Text. *In:* Seale, C. (ed.) *Social Research Methods: A Reader.* Routledge.
Gilbert, M. Social Justice and Web Science: Theory, Praxis and Everyday Life. Web Science 2010, 2010 Raleigh, USA.
Gill, M. C., Zeithaml, V.A. 1985. The Elderly Consumer and Adoption of Technologies. *The Journal of Consumer Research,* 12**,** 353-357.
Goffman, E. 1959. *The Presentation of Self in Everyday Life*, Penguin.
Golbeck, J. 2005. Web-Based Social Networks: A Survey and Future Directions University of Maryland.
Gostin, L. 1997. Health Care Information and the Protection of Personal Privacy: Ethical and Legal Considerations. *Annals of Internal Medicine,* 127**,** 683-690.
Gould, J. D., Lewis, C. 1985. Designing for usability: key principles and what designers think. *CACM.*
Grudin, J., Pruitt, J. Personas, participatory design and product development: An infrastructure for engagement. Participatory Design Conference, 2002 Palo Alto. 144-161.
Harper, R., Rodden, T., Rogers, Y., Sellen, A. 2008. Being Human: Human–Computer Interaction in the Year 2020.

Harrison, S., Tater, D., Sengers, P. 2007. The Three Paradigms of HCI. *Alt. Chi. Session at the SIGCHI Conference on Human Factors in Computing Systems.* San Jose, California, USA.
Hart, J., Ridley, C., Taher, F., Sas, C., Dix, A. Exploring the Facebook Experience: A New Approach to Usability. NordiCHI, 2008. ACM, 471-474.
Hassenzahl, M. 2003. The Thing and I. *In:* Blythe, M., Overbeeke, K., Monk, A.F., Wright, P.C. (eds.) *Funology.* Springer.
Hassenzahl, M., Tractinsky, N. 2006. User experience - a research agenda. *Behaviour and Infomation Technology,* 25**,** 91-97.
Hertzum, M. 2003. Making use of scenarios: a field study of conceptual design. *International Journal of Human-Computer Studies,* 58**,** 215-239.
Hewett, T. T. 2005. Informing the design of computer-based environments to support creativity. *International Journal of Human-Computer Studies,* 63**,** 383-409.
Heyer, C., Brereton, M., Viller, S. Cross-channel mobile social software: An empirical study. CHI, 2008. 1525-1534.
Hirsch, T., Forlizzi, J., Hyder, E., Goetz, J., Stroback, J., Kurtz, C. 2000. The ELDer Project: Social, Emotional, and Environmental Factors in the Design of Eldercare Technologies. *ACM Conference on Universal Usability.* Arlington, Virginia, United States: ACM.
Hooper, C., Frazer, A., Prince, R. 2010a. Can educational games ever be fun? On redesigning the gaming experience. *Interface 2010.* Warwick, UK.
Hooper, C., Millard, D.E. 2010b. Teasing Apart and Piecing Together: Towards Understanding Web-based Interactions *Web Science 2010.* Raleigh, USA.
Hooper, C. a. M., D. E. 2009. Creative Software Engineering (poster contribution). *Grace Hopper Celebration of Women in Computing.* Tuscon, AZ, USA.
Hooper, C. J., Millard, D. E. 2010c. Creative Software Engineering (poster contribution). *"Models of Creative Design" summer school.* Aveiro University, Portugal.
Hooper, C. J., Millard, D. E. 2010d. Supporting Transitions with Creative Software Engineering. *create10.* Edinburgh, UK: eWIC.
Hooper Owens, C., Millard, D. E., Stanford-Clark, A. 2009a. Tugging at the Seams: Understanding the Fabric of Social Sites. *Grace Hopper Celebration of Women in Computing 2009.*
Hooper Owens, C. M., D. E., Stanford-Clark, A. 2009b. Towards a Suite of Tools for Re-imagining Online Social Phenomena. *Interface 2009.* Southampton, UK.
Horváth, I. 2007. Comparison of Three Methodological Approaches of Design Research. *International Conference on Engineering Design.* Paris, France.
Horváth, I. 2008. Differences between ‘research in design context’ and ‘design inclusive research’ in the domain of industrial design engineering. *2008,* 7**,** 61-83.
Howard, S., Vetere, F., Gibbs, M., Kjeldskov, J., Pedell, S., Mecoles, K., Bunyan, M. Murphy, J. 2004. Mediating intimacy: digital kisses and cut and paste hugs. *BCSHCI2004.* Leeds, UK.
Hughes, J., O'Brien, J. Rodden, T. 1998. Understanding Technology in Domestic Environments: Lessons for Cooperative Buildings. *Cooperative Buildings, Integrating Information, Organisation and Architecture.*
Hughes, J., O'Brien, J., Rodden, T., Rouncefield, M. Viller, S. 2000. Patterns of Home Life: Informing Design For Domestic Environments. *Personal technologies,* 4**,** 25-38.

Humphreys, L. 2007. Mobile Social Networks and Social Practice: A Case Study of Dodgeball. *Journal of Computer-Mediated Communication,* 13**,** 341-360.

Introna, L. D., Pouloudi, A. 1999. Privacy in the Information Age: Stakeholders, Interests and Values. *Journal of Business Ethics,* 22**,** 27-38.

Isomursu, M., Tahti, M., Vainamo, S., Kuutti, K. 2007. Experimental evaluation of five methods for collecting emotions in field settings with mobile applications. *International Journal of Human-Computer Studies,* 65**,** 404-418.

Jacko, J. A., Scott, I. U., Sainfort, F., Barnard, L., Edwards, P. J., Emery, V. K., Kongnakorn, T., Moloney, K. P. Zorich, B. S. Older adults and visual impairment: what do exposure times and accuracy tell us about performance gains associated with multimodal feedback? Proceedings of the SIGCHI conference on Human factors in computing systems, 2003.

Java, A., Xiaodan, S., Finin, T., Tseng, B. Why we Twitter: Understanding Microblogging Usage and Communities. WebKDD and SNA-KDD workshop on Web mining and social network analysis, 2007 San Jose, California. 56-65.

John, B. E. Learning and Using the Cognitive WalkthroughMethod: A Case Study Approach. CHI, 1995.

Johnson, G. I., Briggs, P. 1994. Question-asking and verbal protocol techniques. *In:* Cassell, C., Symon, G. (eds.) *Qualitative Methods in Organizational Research: A Practical Guide.* London: Sage.

Johnston, R. B. 2001. Situated Action, Structuration and Actor-Network Theory: An Integrative Theoretical Perspective. *Global Co-Operation in the New Millennium.* Slovenia.

Joyce, M. 1991. Storyspace as a hypertext system for writers and readers of varying ability. *Hypertext '91.* San Antonio.

Kamioka, E. a. Y., S. 2004. Environment-adaptive personal communications realizing ubiquitous computing networks. *Electronics and Communications in Japan,* 1**,** 34-47.

Kazmer, M., Xie, B. 2008. Qualitative Interviewing in Internet Studies: Playing with the media, playing with the method. *Information, Communication and Society,* 11**,** 257 - 278.

Keen, P. 1991. Relevance and Rigor in Information Systems Research: Improving Quality, Confidence, Cohesion, and Impact. *In:* Nissen, H. E., Klein, H.K., Hirschheim, R. (eds.) *Information Systems Research: Contemporary Approaches and Emergent Traditions.* Amsterdam: Elsevier.

Keller, I., Van Der Hoog, W., Stappers, P.J. 2004. Gust of Me: Reconnecting Mother and Son. *Pervasive computing.*

Kelly, H., Et Al 2007. How to build serious games. *COMMUNICATIONS OF THE ACM,* 50**,** 44-49.

Keyani, P., Hsieh, G., Mutlu, B., Easterday, M., Forlizzi, J. 2005. DanceAlong: Supporting Positive Social Exchange and Exercise for the Elderly Through Dance. *CHI.* Portland, Oregon, USA.

Kidd, C. D., Orr, R., Abowd, G. D., Atkeson, C. G., Essa, I. A., Macintyre, B., Mynatt, E., Starner, T. E. Newstetter, W. The Aware Home: A Living Laboratory for Ubiquitous Computing Research. Proceedings of the Second International Workshop on Cooperative Buildings, Integrating Information, Organization, and Architecture, 1999. 191-198.

Kitchenham, B. A. 1996. Evaluating Software Engineering Methods and Tools: Part 1: The Evaluation Context and Evaluation Methods. *ACM SIGSOFT Software Engineering Notes,* 21**,** 11-14.

Kitchenham, B. A., Pickard, L.M. 1998. Evaluating Software Engineering Methods and Tools: Part 9: Quantitative Case Study Methodology. *ACM SIGSOFT Software Engineering Notes,* 23**,** 24-26.

Korhonen, I., Bardram, J.E. 2004. Guest editorial introduction to the special section on pervasive healthcare. *IEEE Transactions on Information Technology in Biomedicine,* 8**,** 229-234.

Laio, C., Tang, H., Liu, Q., Chiu, P., Chen, F. 2010. FACT: Fine-grained Cross-media Interaction with Documents via a Portable Hybrid Paper-Laptop Interface. *ACM Multimedia.* Firenze, Italy: ACM.

Latour, B. 1987. *Science in Action: How to Follow Scientists and Engineerings Through Society*, Open University Press.

Law, E. L.-C., van Schaik, P. 2010. Modelling user experience – An agenda for research and practice. *Interacting with Computers,* 22**,** 313 - 322.

Lei, H., Ranganathan, A. 2004. Context-Aware Unified Communication. *In IEEE International Conference on Mobile Data Management.*

Leichtenstern, K., Erdmann, D., Andre, E. 2008. EVAL - An Evaluation Component for Mobile Interfaces. *MobileHCI.*

Lewis, C., Polson, P. G., Wharton, C., Rieman, J. 1990. Testing a walkthrough methodology for theory-based design of walk-up-and-use interfaces. *CHI.* New York, USA: ACM Press.

Lewthwaite, S. Disability, the Academy and Identity 2.0: Mapping the social experiences of disabled students online. 4th International Conference on e-Social Science, 2008 Machester, UK.

Liscano, R., Impey, R., Yu, Q. Abu-Hakima, S. Integrating Multimodal Messages Across Heterogeneous Networks. First IEEE Enterprise Networking Mini-Conference, 1997. 45-53.

Little, L., Briggs, P. 2009. Pervasive healthcare: the elderly perspective. *PETRA.* Corfu, Greece.

Ljungblad, S., Skog, T., Holmquist, L.E. 2003. From Usable to Enjoyable Information Displays. *In:* Blythe, M., Overbeeke, K., Monk, A. F., Wright, P. C. (eds.) *Funology.* Springer.

Lorincz, K. M., D.J. Fulford-Jones, T.R.F. Nawoj, A. Clavel, A. Shnayder, V. Mainland, G. Welsh, M. Moulton, S., 2004. Sensor Networks for Emergency Response: Challenges and Opportunities. *Pervasive Computing, IEEE,* 3**,** 16-23.

Madsen, K. H. 1994. A Guide to Metaphorical Design. *CACM,* 37**,** 57-62.

Mahlke, S. 2005. Understanding users' experience of interaction. *EACE European association of cognitive ergonomics* ACM.

Mahlke, S., Lindgaard, G. 2007. Emotional Experiences and Quality Perceptions of Interactive Products. *Human-Computer Interaction*.

Maiden, N., Gizikis, A., Robertson, S. 2004. Provoking Creativity: Imagine What Your Requirements Could be Like. *IEEE Software,* 21**,** 68-75.

Maniatis, P., Roussopoulos, M., Swierk, E., Lai, K., Appenzeller, G., Zhao, X., Baker, M. 1999. The Mobile People Architecture. *Mobile Computing and Communications Review,* 1**,** 1-7.

Marti, S. a. S., C. 2001. Active Messenger: filtering and delivery in a heterogeneous network. *Human Computer Interaction,* 20**,** 163-194.

Mather, M., Carstensen, L. 2005. Aging and motivated cognition: the positivity effect in attention and memory. *TRENDS in Cognitive Sciences,* 9**,** 496-502.

Mayo, E. 1933. *The Human Problems of an Industrial Civilization,* New York, MacMillan.

Minocha, S. 2009. A case study-based investigation of students' experiences with social software tools. *New Review of Hypermedia and Multimedia,* 15**,** 245-265.

Morris, M. E. 2005. Social networks as health feedback displays. *IEEE Internet Computing,* 9**,** 29-37.

Mynatt, E. D., Rowan, J., Craighill, S., Jacobs, A. 2001. Digital family portraits: supporting peace of mind for extended family members. *SIGCHI conference on Human factors in computing systems.* Seattle, Washington, United States: ACM.

Nack, F. Capturing Experience - a matter of contextualising events. Proceedings of the 2003 ACM SIGMM workshop on Experiential telepresence, 2003 Berkeley, California 53-64.

Nahm, E., Resnick, B. 2001. Homebound Older Adults' Experiences With the Internet and E-mail. *Computers in Nursing,* 19**,** 257-263.

Nakanishi, Y., Takahashi K, Tsuji T, Hakozaki K 2002. iCAMS: A mobile communication tool using location and schedule information. *Pervasive.* Springer-Verlag.

Namazi, K. H., Mcclintic, M. 2003. Computer use among elderly persons in long-term care facilities. *Educational Gerontology,* 29**,** 535 - 550.

Nanard, J., Nanard, M. 1995. Hypertext Design Environments and the Hypertext Design Process. *CACM,* 38**,** 49-56.

Neable, D. 2009. *RE: Email conversation about software design tools.*

Newell, A. F., Gregor, P. 2004. Design for older and disabled people - where do we go from here? *Universal Access in the Information Society,* 2**,** 3-7.

Nielsen, J. 1988. Hypertext '87. *ACM SIGCHI Bulletin,* 19**,** 27-35.

Nielson, J. 1994. *Ten Usability Heuristics* [Online]. useit.com. Available: www.useit.com/papers/heuristic/heuristic_list.html 2009].

Norman, D. 2003. *Emotional Design: Why we Love (or Hate) Everyday Things*, Basic Books.

Osborn, A. F. 1963. *Applied Imagination: Principles and Procedures of Creative Thinking,* New York, Scribeners and Sons.

Owens Hooper, C., Millard, D., Stanford-Clark, A. Methods for Re-imagining Social Tools in New Contexts (poster contribution). WebSci'09, 2009a Athens, Greece.

Owens Hooper, C., Millard, D., Stanford-Clark, A. 2009b. Towards a Social Fabric for Pervasive Assistive Environments. *PETRA.* Corfu, Greece.

Owens Hooper, C. J., Millard, D.E., Stanford-Clark, A. 2009c. Tugging at the Seams: Understanding the Fabric of Social Sites (poster contribution). *Hypertext.* Torino, Italy: ACM.

Panke, S., Gaiser, B., Werner, B. 2007. Evaluation as Impetus for Innovations in E-learning – Applying Personas to the Design of Community Functions. *Journal of Online Learning and Teaching,* 3.

Paulos, E. 2003. Mobile Play: Blogging, Tagging, and Messaging. *UbiComp.*

Perry, M., Dowdall, A., Lines, K. Hone, K. 2004. Multimodal and Ubiquitous Computing Systems: Supporting Independent-Living Older Users. *IEEE Transactions on Information Technology in Biomedicine,* 8**,** 258-270.

Pfeil, U. 2007. Online social support for older people. *SIGACCESS Newsletter.*

Picard, R. W. 1995. Affective computing. MIT.

Plaisant, C., Clamage, A., Hutchinson, H.B., Bederson, B.B., Druin, A. 2006. Shared Family Calendars: Promoting Symmetry and Accessibility. *ACM Transactions on Computer-Human Interaction,* 13**,** 313–346.

Raman, B., Katz, R. H., Joseph, A. D. 2000. Universal Inbox: Providing Extensible Personal Mobility and Service Mobility in an Integrated Communication Network. *Workshop on Mobile Computing Systems and Applications (WMSCA'00).*

Ramchurn, S. D., Deitch, B., Thompson, M. K., de Roure, D. C., Jennings, N. R., Luck, M. Minimising intrusiveness in pervasive computing environments using multi-agent negotiation. In Proceedings of the 1st International Conference on Mobile and Ubiquitous Systems, 2004.

Reali, P. 2010. *H2 Solve Wicked Problems*, lulu.com.

Richter, A. a. K., M. 2008. Functions of Social Networking Services. *COOP '08 : the 8th International Conference on the Design of Cooperative Systems.*

Ridings, C., Gefen, D. 2004. Virtual Community Attraction: Why People Hang Out Online *Journal of Computer-Mediated Communication,* 10.

Rodden, T., Benford, S. The evolution of buildings and implications for the design of ubiquitous domestic environments. Proceedings of the conference on Human factors in computing systems, 2003. ACM Press, 9-16.

Rodden, T., Crabtree, A. 2004. Domestic Routines and Design for the Home. *Computer Supported Cooperative Work,* 13**,** 191-220.

Ross, J. M. 2010. Informatics Creativity: a Role for Abductive Reasoning? *CACM,* 53**,** 144 - 148.

Roto, V., Ketola, P., Huotari, S. 2008. User Experience Evaluation in Nokia. *CHI.*

Sas, C., Dix, A., Hart. J., Su, R. 2009. Emotional Experience on Facebook Site. *CHI 2009.*

Schuler, D., Namioka, A. 1993. *Participatory Design: Principles and Practices*, CRC Press.

Sellen, A., Eardley, R., Izadi, S., Harper, R. 2006. The Whereabouts Clock: Early Testing of a Situated Awareness Device. *CHI 2006.* Montreal: ACM.

Sengers, P. 2003. The Engineering of Experience. *In:* Blythe, M., Overbeeke, K., Monk, A.F., Wright, P.C. (eds.) *Funology.* Springer.

Sengers, P., Boehner, K., David, Sh., Kaye, J. 2005. Reflective Design. Critical computing: between sense and sensibility. Arhus, Denmark: ACM.

Sengers, P., Gaver, B. 2006. Staying Open to Interpretation: Engaging Multiple Meanings in Design and Evaluation. *DIS.*

Shen, X., Eades, P., Hong, S., Vande Moere, A. 2007. Intrusive and Non-intrusive Evaluation of Ambient Displays. *Pervasive 2007.* Ontario, Canada.

Shneiderman, B. 2000. Creating Creativity: User Interfaces for Supporting Innovation. *ACM Transactions on Computer-Human Interaction,* 7**,** 114-138.

Shneiderman, B. 2007. Creativity Support Tools: Accelerating Discovery and Innovation. *Communications of the ACM,* 50**,** 20-32.

Shneiderman, B., Plaisant, C. 2006. Strategies for Evaluating Information Visualization Tools:

Multi-dimensional In-depth Long-term Case Studies. *BELIV 2006.* Venice, Italy: ACM.

Shock, N. W. 1951. Gerontology (Later Maturity). *Annual Review of Psychology,* 2**,** 353-370.

Silva, P. A. 2009. *Designing User Interfaces with the BadIdeas Method: Towards Creativity and Innovation.* Doctor in Philosophy, Lancaster University.

Society, T. R. 2006. Digital healthcare: the impact of information and communication technologies on health and healthcare.
Spencer, R. 2000. The streamlined cognitive walkthrough method, working around social constraints encountered in a software development company. *CHI.* New York, USA: ACM Press.
Stanford, V. 2002. Using Pervasive Computing to Deliver Elder Care. *IEEE Pervasive Computing,* 1**,** 10-13.
Stone, A. 2003. The Dark Side of Pervasive Computing. *IEEE Pervasive Computing,* 2**,** 4 - 8.
Straker, D., Rawlinson, G. 2002. The 7-Step TRIZ Process. *How to Invent (Almost) Anything.* London: Spiro Press.
Strauss, A. L., Corbin, J. 2004. Open Coding. *In:* Seale, C. (ed.) *Social Research Methods: A Reader.* Routledge.
Suchman, L. 1994. *The Problem of Human Machine Communication*, Cambridge University Press.
Summers, D. 1987. *Longman Dictionary of Contemporary Englishy,* Harlow, Essex, England, Pearson Longman.
Tedjasaputra, A., Sari, E., Strom, G. 2004. Sharing and Learning through Pair Writing of Scenarios. *NordiCHI.* Tampere, Finland.
Terry, M., Mynatt, E. 2002. Recognizing Creative Needs in User Interface Design. *C&C.* Loughborough, UK: ACM.
Ting, T. C. Privacy and Confidentiality in Healthcare Delivery Information System. Computer-Based Medical Systems, 1000 1999 Stamford, CT, USA. 2-4.
Titchener, C. 2005. *Reviewing the Arts*, Lawrence Erlbaum Associates.
Troscianko, E. 2010. *RE: Email exchange about statistical methods. (Personal communication)*
Turk, M. 2005. Multimodal Human Computer Interaction. *In:* Kisacanin, B., Pavlovic, V., Huang, T. S. (eds.) *Real-time vision for human-computer interaction.* Springer.
van Harmelen, M. E. 2001. Interactive System Design Using Oohci Methods. *Object Modelling and User Interface Design: Designing Interactive Systems.* Addison Wesley.
Varshney, U. 2003. Pervasive Healthcare. *IEEE Computer,* 36**,** 138 - 140.
Venkatesh, V., Morris, M. G., Davis, G. B., Davis, F. D. 2003. User acceptance of information technology: toward a unified view. *MIS Quarterly,* 27**,** 425-478.
Wallas, G. 1926. *The Art of Thought,* New York, Harcourt Brace.
Wang, H. J., Raman, B., Chuah, C. Et Al. 2000. ICEBERG: An Internet-core Network Architecture for Integrated Communications. *IEEE Personal Communications (Special Issue on IP-based Mobile Telecommunication Networks).*
Warr, A., O'neill, E. Understanding Design as a Social Creative Process. Creativity and Cognition, 2005.
Warr, A. M. 2007. *Understanding and Supporting Creativity in Design.* Doctor in Philosophy, University of Bath.
Waterworth, J., Ballesteros, S., Peter, C., Bieberm, G., Kreiner, A., Wiratanaya, A., Polymenakos, L., Wanche-Politis, S., Capobianco, M., Etxeberria, I., Lundholm, L. 2009. Ageing in a Networked Society – Social Inclusion and Mental Stimulation. *PETRA.* Corfu, Greece.
Weiser, M. 1989. The computer for the 21st century. *ACM SIGMOBILE Mobile Computing and Communications Review,* 3**,** 3 - 11.

Weiser, M. 1993. Some computer science issues in ubiquitous computing. *CACM,* 36**,** 75-84.
Weiss, R. S. 1995. Preparation for Interviewing. *Learning from Strangers: the art and method of qualitative interview studies.*
Wharton, C., Rieman, J., Lewis, C., Polson, P. 1994. The Cognitive Walkthrough Method: A Practitioner's Guide. *In:* Nielsen, J., Mack, R. (eds.) *Usability Inspection Methods.* John Wiley & Sons.
Wiley, J., Sung, J., Abowd, G. 2006. The Message Center: Enhancing Elder Communication. *CHI.* Montréal, Québec, Canada.
Wright, K. 2002. Social support within an on-line cancer community: an assessment of emotional support, perceptions of advantages and disadvantages, and motives for using the community from a communication perspective. *Journal of Applied Communication Research,* 30**,** 195-209.
Wright, P., Blythe, M., Mccarthy, J. 2006. User Experience and the Idea of Design in HCI. *In:* Gilroy, S. W., Harrison, M.D. (eds.) *Interactive Systems: Design, Specification and Verification.* New York: Springer.
Wright, P., Mccarthy, J. 2008. Empathy and Experience in HCI. *CHI.* Florence, Italy: ACM.
Wright, P., Mccarthy, J., & Meekison, L. 2003. Making sense of experience. *In:* Blythe, M., Overbeeke, K., Monk, A.F., Wright, P.C. (eds.) *Funology: From Usability to Enjoyment.* Kluwer Academic Publisher.
Yin, R. K. 2008. *Case Study Research: Design and Methods*, Sage Publications.

Appendix A: Code Changes to Improve the Social Infrastructure

Figure A-1 shows a class diagram of the prototype social infrastructure. It is followed by information on how and where the functionality for each improvement listed in Section 2.4.4 would be implemented.

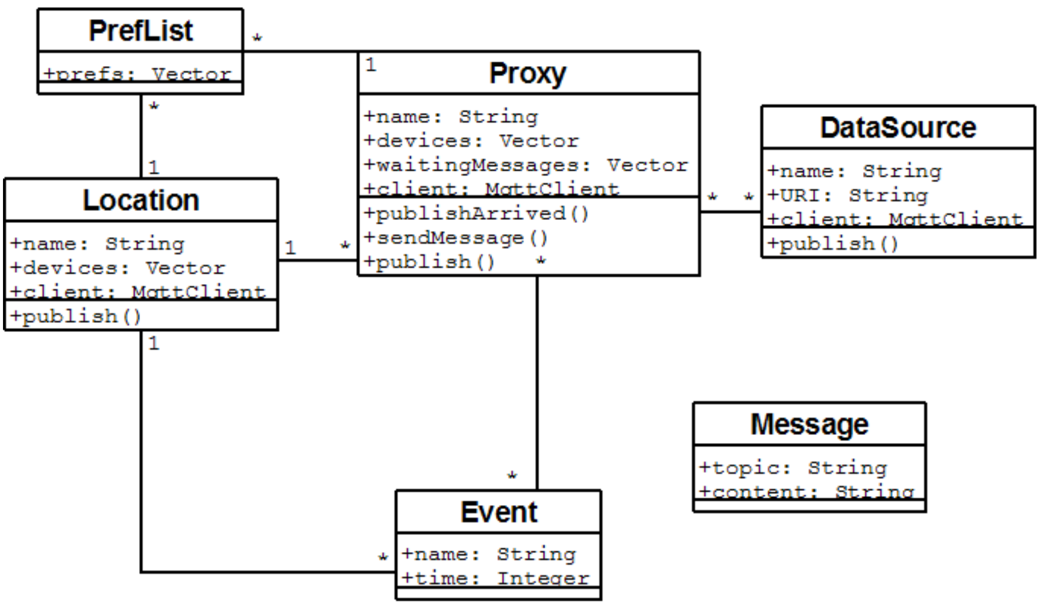


Figure A-1 Class diagram

1. **Proxies may access computers associated with an end user, in order to find out whether the user is active at a computer**

If location is unknown and the recipient is not at an event, check any computers associated with them to see if a computer is active (with this user logged in). If so, send message to the computer, if not, consider whether the person may be en route to a location instead (see below). Implemented in `Proxy.sendMessage`.

Build a '`Device`' class. Each device has a URI, such that the `Proxy` class can interrogate that URI for a response (and find out when the device was last active).

2. **Proxies need awareness of messages' sensitivity and whether output devices are public (e.g. a communal LCD in a corridor is public; a mobile phone is not)**

Augment the `DataSource` class with a 'private' Boolean. As default this is 'true' and any message from that topic is private. If a user marks a `DataSource` as 'public', then incoming messages from that `DataSource` may be received on public devices as well as private ones.

Augment the `Device` class with a 'public' Boolean, which marks whether it is publicly viewable. By default, devices at a location are public, and carried by a person are private.

Enhance the `Proxy.SendMessage` method such that when a suitable available device is found, it is only used if the privacy of the device and the privacy of the message are compatible (that is, it is not a private message and a public device).

3. **Allow a person's subscriptions and messages to be marked as 'high priority'**

Build a `Subscription` class, which holds the identifier of the topic in question and also a Boolean. If the Boolean is true, all items within the subscription are high priority. The `Proxy.subscriptions` vector will hold instances of this class, instead of the identifiers of topics.

Augment `Message` class with a Boolean. If it is true, the message is high priority.

Extend the `Proxy.sendMessage` logic such that high priority messages are (if possible) sent to devices which means they will be seen sooner (e.g. mobile phone, not email inbox).

Extend the `Device` class such that a device can be marked as suitable for high priority messages or not. (Default setting is 'not suitable'.)

4. **Proxies are aware of the relevance of events to non-attendees (e.g. reasoning that Bill's boss is connected with his current meeting)**

Augment event and data source information with ‘content’ tags in a vector: these are keywords to do with the event or topic (e.g. ‘healthcare’, ‘project x’ or ‘social’).

Enhance the `Proxy.sendMessage` logic such that if a message is received during an event, the message is treated as high priority when the keywords associated with the event and the message’s topic match.

5. Proxies reason that just before or after an event, an attendee is likely to be in transit to or from that event

Edit the `Proxy.sendMessage` logic such that if the person’s location is unknown and we are within five minutes of the start or end of an event, they are in transit to or from this event, and devices near the event may be of use.

6. Allow output devices to notify a subscriber of message receipt, only displaying the message content when prompted by the recipient

If a device is public (at a location rather than carried by a person), then ensure it behaves appropriately on message arrival (logic in the `Device.messageArrived` method).

Appendix B: Social Sites' Functionality

This appendix details the primary functionality of various social sites, from data gathered in August 2008.

All sites provide a **profile page**, **friends list**, **news feed** and **private messages**. Table B-1 shows further functionality.

Site	**Functionality**								
	Photos	**Groups**	**Public messages**	**Microblog**	**Blog**	**Videos**	**Apps**	**Music**	**Events**
Bebo	x	x	x	x	x	x	x	x	
Facebook	x	x	x	x	x		x		x
Friendster	x	x	x	x	x	x	x		
LinkedIn		x		x					
Match	x								
MySpace	x	x	x	x	x	x		x	
Orkut	x	x	x	x		x	x		
Twitter		x	x	x					
Windows Live Spaces	x	x	x		x				x

Table B-1: Primary functionality of selected social sites.

Some sites included other prominent functionality:

- **Bebo**: authors, Bebo Nation, to do list
- **Facebook**: birthdays, pokes, instant messaging
- **Friendster**: connections (subtypes: Friends, schools, ‘fan of’, bookmarks, groups), schoolmates, featured fan profiles
- **LinkedIn**: friends list is called contacts list, Q&A, jobs
- **Match**: winking, favourites, searches
- **Orkut**: testimonials, lists (bookmarks, hot/crush/ignore list)
- **Twitter**: view all recent public Tweets. Note: public messages are provided via the @name convention, and groups via the #groupname convention.
- **Windows Live Spaces:** files, lists (books, music, blog, movies, custom)

Examples of peripheral functionality include:

- **Bebo**: blogs, sayings
- **Facebook**: pokes, birthdays, people you may know

Appendix C: Social Sites' Profile Data

This appendix details, for each site, the information which can be provided in the site's profile. The below data was gathered in August 2008:

All profiles include a **name**, **profile picture**, **location** and some amount of **free text**: sometimes the free text is guided (e.g. lists of favourite things, political or religious views), but not always (e.g. "About me"). There is a link between profiles and functionality: for example, a person's profile is linked with that person's list of friends. Similarly, profiles link with other functionality, such as photos, groups, and videos. Table C-1**Error! Reference source not found.** shows further options available in profiles:

Site	**Fields for information on...**					
	Education / work	**Age**	**Gender**	**Relationship**	**URL**	**Contact**
Bebo	x	x	x	x		x
Facebook	x	x	x	x	x	x
Friendster	x	x	x	x	x	
LinkedIn	x				x	
Match	x	x	x	x		
MySpace	x	x	x	x		
Orkut	x	x	x	x	x	x
Twitter					x	
Windows Live Spaces	x	x	x	x		

Table C-1 Profile breakdown of selected social sites

In addition to the material in Table C-1, some profiles allowed for further information, listed below:

- **LinkedIn**: specialties
- **Match** gender/age/location of interest, physical features (height, body type, eye colour, hair colour, body art, best feature), star sign, and many more details
- **MySpace**, star sign, ethnicity, whether the user smokes or drinks
- **Orkut**: fans
- **Windows Live Spaces:** nickname, pets

Appendix D: TAPT Applied to Social Networking

This appendix presents TAPT analyses of aspects of social websites, given in the sequence described in Table D-1.

Functionality	**Site**		
	Facebook	**Orkut**	**Twitter**
Public messaging	1	2	3
Microblogs	4	5	6
Photo-sharing	7		N/A
Groups	8	9	10

Table D-1 Order in which the analyses are presented

1. Public Messaging on Facebook (the 'Wall')

As described, social sites often provide a 'wall', 'whiteboard' or 'scrapbook', where friends and the profile-holder can leave notes. Variations include the ability to augment these notes with HTML formatting or images.

Surface elements

- a box for up to 1000 characters of plaintext[25]
- a 'share' button
- a list of previous messages and activity by the person whose wall this is, most recent first

Literal experienced effects

- quick and easy
- communicating (one to one)
- being overheard

[25] Note: now, in late 2010, Facebook enables the posting of links, images and videos on the 'Wall'. However, at the time of this analysis, it did not.

Abstract experienced effects

- social connectedness (conversing, or letting someone know they are in your thoughts)
- anticipation of a response
- uncertainty (will there be a reply? When? Who else will read the message, how will they respond?).

Piecing Together

A rebuilding of this functionality must take into consideration the above elements and effects. An implementation resulting from this analysis might provide a very simple, clean interface for entering public messages and clarity that the message is public (through showing previous messages left by other people, and perhaps on first use a brief explanation of the mechanism).

One way to implement this in a new context might be to install a microphone and speakers on the door of someone's office. Passers-by may press a button to record a message for the office's owner ("Hi Andy! I dropped by to chat about work, but you weren't here. Catch you later!"); the last ten messages are played in a repeating loop. A time limit (30 seconds, perhaps) reflects the character limits in the original medium. This mechanism is audible to people in the office area, just as public messaging on Facebook is visible to connections in the social network.

(Note: the above demonstrates a holistic approach to reconstruction, as described in Section 4.2.6.)

2. Public Messaging on Orkut (the 'scrapbook')

Public messaging on Orkut largely resembles the same on Facebook, although under a different name (the 'scrapbook').

Surface elements

- a box for free HTML text, limited to 1024 characters
- three buttons: 'post scrap', 'preview' and 'add photo'
- a list of previous messages ('scraps'), most recent first

Literal and abstract experienced effects

These are largely the same as those for Facebook, with two minor differences, both literal:

- quick and fairly easy, rather than quick and easy: the presence of three buttons (not one) make this process slightly more complex
- added expressiveness through HTML-rich formatting and the ability to include images

Piecing Together

An appropriate implementation would be rather similar to that suggested for reconstructing this functionality based on Facebook: differences involve people being able to preview their posts, and include images and more complex formatting with these. The inclusion of images can be reflected by making the set-up a display screen as well as audio equipment, such that audio-video recordings are made. Previews can be incorporated by adding a ten-second time period after recording is complete, during which the most recent recording can be deleted by holding the button for several seconds.

3. Public Messaging on Twitter ('@replies')

Unlike Facebook and Orkut, public messaging is carried out from the home page of Twitter, where users view the stream of tweets from users to whom they are subscribed. By contrast, the other two sites offer public messaging on the profile page of the recipient of the message:

Public messaging on Twitter is achieved via use of the @reply mechanism: users type their message, which like a microblog entry is limited to 140 characters, but include the text "@username" to direct the message at the person with that username. For example, "@bill Are we still on for tea at 10?"

Surface elements

- a plaintext box for up to 140 characters of free text
- one 'update' button
- a list of previous messages, most recent first
- username of the recipient, and the @reply mechanism

Literal and abstract experienced effects

- quick and somewhat easy (additional load: users must understand the @reply mechanism and know the username of the recipient, although they need not navigate to the recipient's profile page in order to send a public message)
- communicating (one to one)
- being overheard

Abstract experienced effects

- social connectedness (conversing, or letting someone know they are in your thoughts)
- anticipation of a response
- uncertainty (will there be a reply? When? Who else will read the message, how will they respond?).

Piecing Together

When considering rebuilding, one must remember that like Facebook and unlike Orkut, public messaging on Twitter involves plain text only: additionally, updates are limited to 140 characters. Obvious changes to the existing reconstruction (the voice recorder on the office door) would therefore be to provide functionality to leave audio-only messages limited to, say, 15 seconds in length. This does not quite capture public messaging on Twitter, however, as one would still have to walk to the office door of the recipient, whereas Twitter enables messages to be left from the homepage. Instead, the audio equipment could be in the reception of our office building, and people leaving messages must speak the name of their intended recipient for clarity.

4. Microblogging on Facebook (‘Status Updates’)

Microblogging involves posting very succinct text updates, generally limited to 140 or 160 characters. Java et al (Java, 2007) suggest that the constraint on message size increases the speed of communication. They theorize that the constraint on message length lowers the time and thought investment from a microblogger (as opposed to a traditional blogger); this is reflected by the fact that microbloggers tend to post more frequently than bloggers, perhaps daily rather than weekly.

The presentation of microblogging differs in different sites: for example, it is foregrounded in Twitter (where it is the primary functionality), but mixed with other information in Facebook.

Surface elements

- a box for a limited amount of free plaintext (420 characters)
- a ‘share’ button
- a list of previous microblogs updates (your own, or others’)
- buttons with the option to add URLs, images, videos or application-specific materials

Literal experienced effects

- fairly quick and easy
- communicating (one to many)
- broadcasting information

Abstract experienced effects

- presence in the community – consolidate online identity by adding more data
- openness about current experiences
- anticipation of responses

- uncertainty about responses and audience: especially if privacy settings are low, and anyone can access the content. Even if a very specific group of people can access the content, it is not guaranteed that they will do so, and thus uncertainty remains

Piecing Together

Piecing these elements and effects together in a new environment would again include a very simple design allowing the composition and posting of microblogs. It is important to incorporate clarity about the audience, as controlled by privacy settings such as “friends only”, “friends of friends” or “anyone”.

One reconstruction of this functionality might see the microblogger wearing a t-shirt which incorporates a scrolling text display[26] displaying their most recent post, although this does not capture the ability to incorporate graphical information. (This system would be visible to people in the physical vicinity of the microblogger, rather than to friends on the microblogger’s online social network. Clarity about the microblogger’s audience comes from their own awareness of where they are and who is around them.)

5. Microblogging on Orkut (‘status updates’)

Microblogs on Orkut are presented a little different than on Facebook: there appears to be no history of previous status updates.

Surface elements

- a box for a limited amount of free text (140 characters), which can include emoticons selected from a dropdown list (nine emoticons available)
- an ‘edit’ button (this is also used to make the first, fresh post)

Literal and abstract experienced effects, and Piecing Together

Again, experienced effects are largely similar to those associated with microblogging on Facebook, although updates to the status seem simpler here, as there are fewer formatting options (e.g. no option to include images). As such, a similar rebuilding seems appropriate.

6. Microblogging on Twitter (‘Twittering’)

Twitter is often presented as a platform primarily for microblogging: certainly, the option is foregrounded on this website, which presents a mechanism for posting

[26] Such t-shirts are currently on the market: they use thin battery-powered LED displays to show short messages.

updates at the top of the homepage. It is worth noting that the surface elements of microblogging on Twitter are almost identical to those for public messaging: indeed, public messaging is achieved by using microblogging in a slightly different way (as described in the Teasing Apart of public messaging on Twitter, above).

Surface elements

- a box for a limited amount of free plaintext (140 characters)
- an 'update' button

Literal and abstract experienced effects, and Piecing Together

As with Orkut, experienced effects are very similar to those found with Facebook, although again, the act of posting an update is even quicker and easier than either of the previous two platforms, as there are no options to add URLs, videos, emoticons or similar[27]. Again, therefore, a similar reconstruction seems appropriate.

7. Photo-sharing on Facebook and Orkut

The majority of the sites surveyed allowed users to upload and caption photos, which can be commented upon by the photo's owner or other users. Some sites, including Facebook and Orkut, allow users to 'tag' friends in photos, adding metadata which links images with people's profiles. Both Facebook and Orkut allow photo upload, captioning and tagging. Indeed, it would appear that surface elements and experience effects are nearly identical between the two sites, as both offer basically the same functionality.

Surface elements

- a (generally complex) photo upload process
- the option to annotate images with text
- the option to 'tag' images (indicate which contacts are displayed)
- the ability to view photos

Literal experienced effects

- broadcasting information
- sharing past experiences

Abstract experienced effects

- presence in the community – consolidate online identity by adding more data
- openness about past experiences

[27] Again, note that the above analysis was conducted in 2008. Nowadays, third party apps such as Twitvid and Twiddeo enable the sharing video material via Twitter.

- anticipation of discussion about these experiences
- reminiscence
- uncertainty about responses and (depending on privacy settings) audience

Piecing Together

Again, Piecing Together should incorporate the above items. There must be a method for placing photos (and annotations) in the shared space, perhaps involving photo selection via an appropriate interface (e.g. computer monitor; TV screen; digital photo frame). Users must be able to browse and annotate their online photos, and be notified if people comment. These functionalities combine to allow users to feel that they are sharing memories.

One re-envisioning might involve a novel tabletop that displays a sequence of photographs uploaded by the table's owner.

8. Groups on Facebook ('Groups')

Online groups enable people with shared interests to maintain contact. It is worth noting that groups (also described as 'communities' or 'forums' – although the word 'forum' can also refer to a specific format of webpage which supports threaded discussions) constitute a wide array of functionality, from browsing and searching to find groups, to viewing groups recently joined by friends, to viewing very active groups. For the purposes of this Teasing Apart, we consider the experience of searching for and joining a group about line dancing.

Surface elements

- a plaintext box to enter search text
- a 'search' button
- after searching, a list of matched results: this list displays for each group a title, icon, number of members and type (e.g. "EntertaiUDent & Arts – Dance"). It also includes recent activity (e.g. "3 More Members, 1 Wall Post"). Each group has a link, allowing the user to either join the group, or request to join subject to admin acceptance.
- after clicking on a group, the group's webpage is shown: this is akin to a person's profile. It contains the above data and a description, contact details, location, discussion board, admins, members, recent news, public message area, photos, links, videos and related groups.

Literal experienced effects

- quick
- subject to suitable search text, easy
- availability of relevant communities

- availability of information about these communities – e.g. popularity (based on membership numbers) and activeness (based on recent activity)
- ability to join these communities

Abstract experienced effects

- potential connectedness and online presence

Piecing Together

Many hobby groups and communities exist in the physical world, but the process of searching for these is less simple than online. One analogy to the above is browsing through listings in a local newspaper: however, this does not include common classifiers of the groups on offer, nor information on recent activity – not to mention that the activity is browsing, not searching.

An appropriate rebuilding might involve building a database of information on local community groups, and enabling search-based access to this via a number of interfaces, including the web and smart phones. Recent information on shifts in membership would be difficult to maintain, as it would require regular updates from each group: an alternative way to gauge group popularity would be to display the number of people who have looked up that group in the last month (or to allow a ratings system, like that used for sellers on eBay), while activity can be conveyed by including information on how frequently the groups meet. Similar descriptors to those above (descriptions, contact details, photos etc) can be stored in the database, and made available to searchers.

9. Groups on Orkut ('Communities')

Searching for groups on Orkut is not dissimilar to the same experience on Facebook, but differences do exist. Differences are shown below:

Surface elements

- after searching, a list of matched results, first showing groups in the user's country, then showing groups worldwide. For each group, the list displays a title, icon, category, location, number of members and brief description.
- After clicking on a group, the group's webpage is shown. This has the above data and language, owner, type (public, moderated, validated), content privacy (open or closed to non-members) and creation date. If they exist, the group's forums (with topics, number of posts, and date of last post), polls and events are displayed.

Literal and abstract experienced effects

These are largely the same as those for Facebook, with minor differences:

- information about how active groups are is less obvious (but exists upon clicking on a search result)
- joining is slightly harder (one must click on a community from the search results in order to then join it)

Piecing together

As such, a Piecing Together of the experience of searching groups in Orkut would be largely similar to reconstructing the same action in Facebook.

10. Groups on Twitter ('#hashtags')

As with public messaging, the functionality associated with implementing groups is different on Twitter to Facebook and Orkut. Twitter uses hashtags, inline to microblogs posts, by prefixing words with the has symbol. For example: #interest. Example uses suggested by Twitter are to collate Tweets relating to

- events or conferences (e.g. "Tara's presentation on communities was great! #barcampblock")
- disasters ("#sandiegofire A shelter has opened up downtown for fire refugees.")
- context ("I can't believe anyone would design software like this! #microsoftoffice")
- recall ("Buy some toilet paper. #todo"), and
- quotes ("Great minds discuss ideas. Average minds discuss events. Small minds discuss people. ~Eleanor Roosevelt #quote").

As such, groups on Twitter work somewhat differently to groups on Facebook or Orkut, primarily adding context and metadata, like tags, enabling like-tagged posts to be collated. By contrast, groups on 'heavier' social sites have a specific membership of users, along with functionality such as public messaging, photo-sharing and so on.

It is worth noting that it is quite difficult to find documentation of the hashtag functionality on Twitter, without obvious links on the main help pages. However, at least two web pages exist from which hashtag groupings can be browsed or searched: http://hashtags.org/ and http://twittgroups.com/[28]. Both sites offer search experiences, as follow:

Surface elements

- a plainbox to enter search text

---

[28] Last accessed October 2010.

- a ‘search’ button
- knowledge of the #hashtag mechanism
- after searching using hashtags.org, a list of matching tags and messages. Clicking on matching tags shows all messages with this tag.
- (this effect applies not to Twitter but to the augmented functionality provided by TwittGroups.com): after searching using twittgroups.com, a list of groups which match this the search term are shown: each has a title, brief text description and URL. Clicking on a group shows further information (description, icon, etc.; also public messaging and a list of members) and an option to join: clicking on ‘join’ opens a new Twitter window with a pre-written message stating you have joined the group, ready to be submitted via the ‘update’ button.

Literal experienced effects

- quick
- subject to suitable search text and knowledge of the #hashtag mechanism, easy
- availability of relevant communities/activity around this topic
- knowledge of popularity of this topic, based on quantity of results
- ability to join communities

Abstract experienced effects

- potential connectedness and online presence

Note that the Twittgroups website appears to be trying to provide more traditional ‘groups’ in the sense of the interpretation found on sites such as Facebook and Orkut, by augmenting the functionality with items such as public messaging, member lists and so forth. However, functionality provided by Twitter alone is that available via the Hashtags site, and it is this which we should consider for reconstruction.

What makes Twitter groups different from others is that the groups exist only because tagged microblog posts exist: participants need only use the group’s hashtag in a post, rather than join a pre-defined group and then post to it.

Piecing Together

This experience boils down to seeing all recent public posts which are tagged with the search term. A rebuilding must enable people to enter search text and access the relevant information: one approach might be to allow people to type their term into a dedicated keyboard situated beside a microprinter (or select images which represent concrete search terms, such as ‘oak trees’ or ‘family’): the microprinter promptly prints matching tagged Tweets. An augmentation would be a switch which when in one position means searches are confined to hashtags, and when in the other means that searches cover all messages. One can envision a parallel implementation whereby companies have ‘buckets’ of relevant Tweets: for example, an IBM reception might have a container below a printer which dispenses in real-time slips of paper printed with Tweets about the Smarter Planet initiative.

Appendix E: Comparative Evaluation Materials

Participants were given various written materials at the beginning of this experiment. These included the Participant Information Sheet, Consent Form, and an 'Introduction and Plan' which outlined more detailed information about the 2.5 hour session. The text of these documents follows:

Participant Information Sheet

Study Title: Evaluation of the TAPT process
Researcher: Clare Hooper
Ethics number: N/09/09/004

Please read this information carefully before deciding to take part in this research. If you are happy to participate you will be asked to sign a consent form.

What is the research about?

The purpose of this study is to evaluate the Software Engineering method called TAPT, or Teasing Apart, Piecing Together. This is to be achieved by asking participants to carry out similar tasks using either TAPT, traditional software engineering methods, or Unstructured Discussion, and then analysing the results.

Why have I been chosen?

You have been approached to participate in this study because you are a software engineering professional, and as such you are someone who might use the methods being evaluated.

What will happen to me if I take part?

During the study, participants will be asked to work in groups to generate two software designs, and deliver these in written form. The investigator will keep written results, but these will be kept private.

Each participant will be asked to fill in four written questionnaires: one for each of the two software design exercises, plus one opening and one closing questionnaire.

These are intended to elicit participants' experiences of the methods, to help evaluate those methods.

At the end of the session, a group discussion and debrief will be held. This is an opportunity for all participants to explore the value of the methods used, and to ask any questions which have arisen during the session. This discussion will be recorded, but again, information will be kept private and transcripts will be anonymised.

The whole study should take no more than 2.5 hours, and no follow up is planned.

Are there any benefits in my taking part?

By taking part, you have the opportunity to learn a new software engineering design technique. You may also view the study as an opportunity to be refreshed on alternative design techniques (you will be asked to apply an alternative technique), and to meet other participants within IBM. Light refreshments will be provided during the study.

Are there any risks involved?

There is one minimal risk involved in taking part in this study, which is that you may become tired over the 2.5 hour duration. To alleviate this risk, light refreshments are provided. It is important to note that you are welcome to pause for a rest, or fully withdraw from the study, at any time.

Will my participation be confidential?

Yes. Your personal details will not be included on any written materials which you are asked to provide. Audio recordings will be edited to assure anonymity: if, for example, a person's name is mentioned, the stored data will be edited to remove this detail.

What happens if I change my mind?

You have the right to withdraw at any time without your legal rights being affected.

What happens if something goes wrong?

If you have a concern or complaint, please contact the ECS School Office on school@ecs.soton.ac.uk or school@ecs.soton.ac.uk02380 592909.

Where can I get more information?

Feel free to get in touch with Clare Hooper or her supervisor, David Millard:

Clare: clare@ecs.soton.ac.uk, 01962 816863 (x246863)
David: dem@ecs.soton.ac.uk, 023 8059 5567

Who is sponsoring this research?

The research sponsor is the School of Electronics & Computer Science at the University of Southampton. Clare Hooper is a full-time student based jointly in ESC and here at IBM Hursley.

Consent Form

Study title: Evaluation of the TAPT process
Researcher name: Clare Hooper
Study reference: Evaluation of software engineering design processes
Ethics reference: N/09/09/004

Please initial the box(es) if you agree with the statement(s):

I have read and understood the information sheet (November 2009 / Version 1.0) and have had the opportunity to ask questions about the study. ☐

I agree to take part in this research project and agree for my data to be used for the purpose of this study. ☐

I understand my participation is voluntary and I may withdraw at any time without my legal rights being affected. ☐

I agree to carry out two software design exercises, and provide the written outputs of these exercises. I understand that no data which can identify me will be stored. ☐

I agree to fill in four written questionnaires, and I understand that my personal data will not be stored. ☐

I agree to participate in a group discussion and debrief. ☐

I agree to the recording of the group discussion and debrief: I understand that no data which can identify me will be stored. ☐

Name of participant (print name)…………………………………………………

Signature of participant……………………………………………………..

Name of Researcher (print name) ………………………………………………

Signature of Researcher……………………………………………………

Date………………………………………………………………………

Consent Form [November 2009] [Version 1.0]

Introduction and Plan

Thank you for participating in this study, which is designed to evaluate different software engineering approaches to re-imagining experiences in new contexts. Please bear in mind that this study is designed to evaluate the efficacy of the software engineering approaches, and not your ability to apply them – you are not being assessed!

Once you have read this information, please place yourselves in pairs. Each pair will be given a group name, and each person a participant number.

If anything is unclear at any point, please ask for clarification.

Today's plan:

1430: Explanation of the procedure and Q&A
1440: First design exercise, followed by questionnaire
1540: Second design exercise, followed by questionnaire
1640: Discussion and debrief
1700: Close

Feel free to help yourselves to snacks throughout.

Methods to test

Method 1: Scenarios and Personas

Scenarios are a tool for imagining ideal user interactions.

1. Consider what you want to design, and carry out unconstrained and uncritical brainstorming. Given the time limits of this exercise, please be prepared to move on after ten minutes.
2. Consider types of user of the system you're designing (personas). Each persona represents a set of people who will use the system in a distinct way. Create the persona(s) which seem key to the system. An example persona is given below.
   You *may* wish to consider aspects such as: attitudes, experiences, aspirations, and other factors which will influence the persona's expectations; behaviours the persona will expect or desire from the system; how the persona thinks about data in the system.
3. Construct scenarios – an example is given below. When you do this, consider the broad context of use: how does the system help your personas achieve their goals? Don't describe product or interaction detail, but high-level actions from the user's perspective. You might want to consider the setting of use, how long each usage session lasts, and the primary activities during use. It might be helpful to consider requirements of your personas. Construct multiple scenarios, if this is helpful.

Example persona: Vivien Strong, a real-estate agent in Indianapolis, whose goals are to balance work and home life, close the deal, and make each client feel he is her *only* client.

Example scenario: While getting ready for work in the morning, Vivien uses her phone to check her email. It has a large enough screen and quick connection time that it's more convenient than booting up a computer as she rushes to make her daughter, Alice, a sandwich for school.

Vivien sees an email from her newest client, Frank, who wants to see a house this afternoon. The device has his contact info, so now she can call him with a simple action right from the email.

Whilst on the phone to Frank, Vivien switches to speakerphone so she can look at the screen while talking. She looks at her appointments to see when she's free…

Based on chapter 6 of About Face: The Essentials of Interaction Design (Alan Cooper, Robert Reimann, David Cronin, Wiley, 2007.

Method 2: Group Discussion

A very informal method, this simply involves carrying out a verbal discussion within your group about how to solve the problem at hand. Please do not write or draw ideas during your discussion, but write a paragraph describing your envisioned design once you have decided upon its details.

Method 3: Teasing Apart, Piecing Together (TAPT)

Stage 1: Teasing Apart

Stage one of TAPT helps people understand and 'distil' the nature of an experience. The output of 'Teasing Apart' is a table showing how the experience breaks down.

Approach

The following table describes each step of Teasing Apart. The numbers shown in brackets show the order in which to fill in the table.

<table>
<tr><th colspan="5">Description of Teasing Apart, step by step</th></tr>
<tr><td rowspan="3">Experience (1) Brief description of the chosen functionality and the experience of using it.</td><td rowspan="3">Surface elements (2) These are generally nouns ('line', 'box', 'arrangement of photos') and adjectives ('bold', 'simple', 'complex') relating to the design.</td><td colspan="2">Experienced effects These focus on the physical, emotional and intellectual effect upon participants, and tend to be abstract nouns ('excitement'), noun/verb pairs ('hunger sated') and perhaps adverbs ('quickly'). There are two types of effect, shown below…</td><td rowspan="3">Distilled experience (5): Consider your table of information, particularly the aspects which you think are key to the experience, and use it to describe the experience as a sentence. Try to keep your sentence neutral: for example, you might mention 'broadcasting' information rather than 'showing' it, because 'showing' implies a visual broadcast.</td></tr>
<tr><td>Literal (3.1) Concrete results such as a loud noise, 'broadcast information'.</td><td>Abstract (3.2) Relating to emotional and intellectual effects, such as 'excitement', 'co-experience'. This step is important: dig deep!</td></tr>
<tr><td colspan="2">(Step 4) Review the above lists of effects, and identify effects that seem especially important, unique or key to the experience. Underline them.</td></tr>
</table>

The lists of elements and effects which you generate will vary in length according to the experience you are Teasing Apart; so will how many effects you think are key. Don't worry about generating a specific number of items: if no more items occur to you, you're probably done.

The following two tables give examples of the Teasing Apart phase:

<table>
<tr><th colspan="5">Teasing apart pulling a Christmas cracker</th></tr>
<tr><th rowspan="2">Experience (1)</th><th rowspan="2">Surface elements (2)</th><th colspan="2">Experienced effects (3 & 4)</th><th rowspan="2">Distilled experience (5)</th></tr>
<tr><th>Literal (3.1)</th><th>Abstract (3.2)</th></tr>
<tr><td>When two people pull a Christmas cracker, it splits (with a bang) into two uneven parts. A cracker usually contains some small, cheap novelties.</td><td>* Cheap and cheerful: brightly wrapped
* Plastic toy and joke
* Paper hat
* Pulled with a friend</td><td>* Bang (when it works)
* Offered to one another
* Pulled together
* <u>Hiddenness</u> of contents (4)</td><td>* <u>Sharing</u> (4)
* Co-experience
* <u>Cultural connotations</u> (4)
* Excitement
* <u>Suspense</u> (when pulling) (4)
* <u>Surprise</u> (does it go bang? What is inside?) (4)</td><td>A shared experience associated with Christmas. Participants don't know the contents of their shared item: this lack of knowledge leads to suspense and surprise.</td></tr>
</table>

Teasing apart photo-sharing on Facebook				
Experience (1)	**Surface elements (2)**	**Experienced effects (3 & 4)**		**Distilled experience (5)**
		Literal (3.1)	**Abstract (3.2)**	
Facebook allows users to users to upload and caption photos, which can be commented upon by the photo's owner or other users. Viewers can 'tag' friends in photos, adding metadata which links images with people's profiles.	* a somewhat complex photo upload process * the option to annotate images with text * the option to 'tag' images, indicating who is shown * the ability to view photos	* broadcasting visual information (4) * sharing past experiences	* presence in the community – consolidate online identity by adding more data(4) * openness about past experiences(4) * anticipation of discussion about these experiences * reminiscence(4) * uncertainty about responses and (depending on privacy settings) audience	A way to share and annotate imagery from the user's past; their audience can view and annotate that imagery.

TAPT Stage 2, Piecing Together

The second phase of TAPT is a creative tool for generating ideas. It involves re-imagining a teased apart experience in a new context, whether physical or digital. For example, it was first used to capture aspects of the experience of pulling Christmas crackers, enabling the 'cracker experience' to be provided on a website. The output of 'Piecing Together' is a description of the experience in its new context.

Approach

Piecing together involved taking a distilled, teased apart experience, and re-providing it in a new context. There will always be many possible ways to re-provide an experience, so there are no wrong answers: choose what seems right.

The steps of Piecing Together are thus:

4) Brainstorming, particularly using the key effects identified earlier. Feel free to use scrap paper, whiteboards etc. if this is helpful. If you find it hard to come up with ideas, consider things you might change from the original experience. These might include:
    - the modality of communication (vocal, textual, musical, graphical, symbolic)
    - the technology used (pen and paper, PDA, telephone, cardboard)
    - scale (are we working with one person? Tens? Hundreds? How big is the physical space? How large are the items with which people interact?)

   You might also want to think about what technologies traditional occur in the original and the new environments – office might traditionally include telephones, whiteboards and desktop computers, while parks tend to traditionally have benches, maps printed on large boards and fountains. What can you use?

5) Scenario building. Using your ideas, build an example reconstruction. Below are two examples:

Piecing together pulling a Christmas cracker, on the web	
Distilled experience	A shared experience associated with Christmas. Participants don't know the contents of their shared item: this lack of knowledge leads to suspense and surprise.
Context	The web.
Description	A web page with a very simple (but cheery) design and graphics. The virtual cracker contains a web toy (game, illusion etc) and joke, alongside a mask image which can be printed and cut out. The virtual cracker is sent by email – the sender can't see the contents of the cracker until the recipient has 'pulled' it by opening the link. Suspense and surprise are aided by showing no contents on the first page; very slowly changing the page when the cracker is pulled; and playing a .wav file (of a bang), when it works.

Second example of Piecing Together:

Piecing together photosharing on Facebook	
Distilled experience	A way to share and annotate imagery from the user's past; their audience can view and annotate that imagery.
Context	The home.
Description	A novel tabletop which displays a sequence of photographs uploaded by the table's owner. The tabletop incorporates a touchscreen interface to allow the owner to place photos and annotations into the shared space. Any passerby can use the same interface to annotate photographs with commentary.

6) Check your reconstruction:
    i. Considering the distilled experience and the key elements of the original experience, have you included everything you want to? (It may be that you choose to drop some key effects from the original experience. This is fine.)
    ii. Have you incorporated any new key effects which you did not intend?
    iii. Refine the pieced together scenario accordingly.
    iv. Repeat the above three steps until you are happy with the scenario.

For example, the rebuilding of 'photo-sharing' lacks the effect of 'uncertainty' about the audience: this might be *added* by positioning the table in a public space instead of the home, or perhaps altering the photo display so that it is visible from the street outside the home.

Blank form for TAPT results

Teasing apart ______________________________				
Experience (1)	**Surface elements (2)**	**Experienced effects (3 & 4)**		**Distilled experience (5)**
		Literal (3.1)	**Abstract (3.2)**	
		Remember to note key effects!		

Piecing together ______________________________	
Distilled experience	
Context	
Description	

Questionnaires

Opening questionnaire

In this session, you still try out some software design processes, and be asked for your opinion on these. What is your previous experience of software design processes? (Consider, for example, scenario building, storyboarding and UML.) Do you perceive these processes as generally positive or negative?

…………………………………………………………………
…………………………………………………………………
…………………………………………………………………
…………………………………………………………………

What participant number are you?

…………………………………………………………………

Main Questionnaire

What is your participant number?

………………..………………..………

Which task did you work on?

Picnic ☐ Wiki ☐ Social networking ☐

Which method did you use?

Scenarios ☐ Group discussion ☐ TAPT ☐

Clarity of explanation

The following questions are intended to explore how clear the exercise was.

Were you able to carry out the task? If not, why?

………………..………………..………………..……………………………………....

Were the task and the method explained clearly? If not, how so?

………………..………………..………………..……………………………………....

Did you find any steps of the method difficult? If so, which ones, and how so?

………………..………………..………………..……………………………………....

………………..………………..………………..……………………………………....

………………..………………..………………..……………………………………....

Other methods

Do you think there are other methods that you might have used to do this sort of work? (Ignoring any which you have used during this study.) Please explain why they are better or worse than the one you applied.

…………………………………………………………………

…………………………………………………………………

…………………………………………………………………

Understanding and replicating experience

The following table is intended to explore how useful the method you used was in helping you understand and re-create experiences. Please write any thoughts you have in the space provided in the second and third columns; the first column gives guidance on what we're looking for.

Ratings: Please provide a rating in answer to each question. Choose from Very Well (VW), Well (W), Okay (OK), Badly (B), Very Badly (VB).

Aspect of interest	Increased understanding	Replication of original experience
Overall…	How well did the method do at improving your understanding of the original experience? How so? Rating:	How well does your artefact recreate the experience of the original? Does it support experiences analogous to the original?
Emotional aspects. For example, these might include feelings of diligence, happiness or anticipation.	How well did the method do at improving your awareness of emotional effects of the original experience? Why?	How well does your redesigned experience replicate the emotional aspects of the original? Would users experience similar emotional reaction as with the original?
The social context and social effects of the experience. For example, people's social expectations beforehand and social fulfilment afterwards.	How well did the method improve your understanding of the social context of the original experience? Why?	How well does your design replicate the social context of the initial experience? In what way?
User perceptions, expectations and reactions during the sequence of events.	How well did the method improve your understanding of the changing perceptions, expectations and reactions of users before, during and after the experience you analysed? In what way?	How well will user perceptions, expectations and reactions around the new experience match those of the original? Why?

More questions on usefulness

How well did the method scaffold a discussion about the problem?
Very well ☐ Well ☐ Ok ☐ Badly ☐ Very badly ☐
Why? ….……………………………………………………………
…………………………………………………………………
…………………………………………………………………

Some of the approaches in this study were formal (Scenarios and TAPT), and some less so. What was the effect of this structure or lack of structure on the way you worked?
…………………………………………………………………
…………………………………………………………………
…………………………………………………………………

How well did the method support you in creating imaginative and novel designs?
Very well ☐ Well ☐ Ok ☐ Badly ☐ Very badly ☐
Why?…..………………………………………………………
…………………………………………………………………
…………………………………………………………………

How well are the artefacts you generated (written text, tables, concept maps, scenarios) suited to use in the workplace? (I.e., for documentation or communication.)
Very well ☐ Well ☐ Ok ☐ Badly ☐ Very badly ☐

How so?..………………………………………………………
…………………………………………………………………
…………………………………………………………………

How well did you replicate the surface design elements of the original experience? (These are tangible design elements, such as textboxes and graphics in a webpage or shiny wrapping paper and a paper hat in a Christmas cracker.)
Very well ☐ Well ☐ Ok ☐ Badly ☐ Very badly ☐
Why? ……………………………………………………………
…………………………………………………………………
…………………………………………………………………

Considering the method as a whole, do you have any other comments about its usefulness or value?
…………………………………………………………………
…………………………………………………………………
…………………………………………………………………

Closing questionnaire

Considering the methods you have applied today, which did you find easier to use? (Consider how straightforward the methods were, and how long they took to apply.) Why?

……………………..……………………..………………
……………………..……………………..………………
……………………..……………………..………………
……………………..……………………..………………

Which did you feel produced more useful (powerful, practical, relevant) results? Why?

……………………..……………………..………………
……………………..……………………..………………
……………………..……………………..………………
……………………..……………………..………………

Which did you feel was more effective at replicating the experience you were considering? Why?

……………………..……………………..………………
……………………..……………………..………………
……………………..……………………..………………
……………………..……………………..………………

What participant number are you?

…………………………………………………………………

Finally, a big **thank you** for your time and effort. It is very much appreciated!

Appendix F: Numeric Data on Comparative Evaluation

TAPT Artefacts

The following table spans three pages.

	Microblog groups					Picnic groups						Wiki groups					
	B	**I**	**K**	**N**	**P**	**C**	**F**	**I**	**L**	**O**	**Q**	**A**	**D**	**E**	**G**	**J**	**M**
Experience description																	
Sentence count	1	1	1	1	2	3	1	3	2	1	4	3	2	N/A	3	1	3
Word count	15	23	12	20	30	26	12	19	26	15	29	29	25	33	33	32	31
Surface elements																	
Number of elements	10	7	8	8	6	9	7	7	7	9	7	4	4	8	6	5	6
Number of elements repeated by other groups	3	3	4	5	2	8	6	5	7	7	5	3	4	6	5	4	4

	Microblog groups					Picnic groups						Wiki groups					
	B	**I**	**K**	**N**	**P**	**C**	**F**	**I**	**L**	**O**	**Q**	**A**	**D**	**E**	**G**	**J**	**M**
Percentage elements repeated by other groups	30	43	50	63	33	89	86	71	100	78	71	75	100	75	83	80	67
Literal effects																	
Number of effects	4	3	3	4	5	7	10	4	3	8	4	2	3	6	4	6	2
Number of effects repeated by other groups	2	2	3	2	2	2	5	2	3	6	4	0	2	3	2	3	1
Percentage of effects repeated by other groups	50	67	100	50	40	29	50	50	100	75	100	0	67	50	50	50	50
Number of key effects	1	0	3	2	2	5	3	0	2	3	3	0	1	2	2	0	0
Number of key effects repeated by other groups	1	0	2	0	1	2	3	0	2	3	3	0	1	1	1	0	0
Percentage of effects identified by this group as key	25	0	100	50	40	71	30	0	67	38	75	0	33	33	50	0	0
Percentage of key effects identified as key by other groups	100	0	67	0	50	40	100	0	100	100	100	0	100	50	50	0	0
Abstract effects																	
Number of effects	7	4	4	4	6	3	9	5	4	6	5	3	5	4	5	6	4
Number of effects repeated by other groups	2	3	3	2	4	3	5	5	3	3	3	1	3	1	2	3	1

	Microblog groups					Picnic groups						Wiki groups					
	B	**I**	**K**	**N**	**P**	**C**	**F**	**I**	**L**	**O**	**Q**	**A**	**D**	**E**	**G**	**J**	**M**
Percentage of effects repeated by other groups	29	75	75	50	67	100	56	100	75	50	60	33	60	25	40	50	25
Number of key effects	5	0	2	1	2	2	4	0	3	3	3	1	3	3	2	5	2
Number of key effects repeated by other groups	2	0	1	0	2	2	3	0	3	1	2	0	2	1	2	2	0
Percentage of effects identified by this group as key	71	0	50	25	33	67	44	0	75	50	60	33	60	75	40	83	50
Percentage of key effects identified as key by other groups	40	0	50	0	100	100	75	0	100	33	67	0	67	33	100	40	0
Distilled experience																	
Sentence count	1	3	1	1	2	3	1	3	1	1	1	1	2	1	3	2	1
Word count	11	36	16	19	20	32	20	29	20	11	18	21	17	10	31	28	13
Context of new exp																	
Word count	3	2	2	3	7	1	2	1	1	2	1	5	5	3	22	14	9
New experience																	
Sentence count	2	3	3	10	N/A	10	5	5	7	4	2	7	3	N/A	6	5	4
Word count	39	44	47	78	31	102	62	64	67	41	56	59	33	98	76	29	51

Table F-1 Numeric per group data on TAPT artefacts

	Microblog		Picnic		Wiki	
	Average	SD	Average	SD	Average	SD
Experience description						
Sentence count	1.20	0.45	2.33	1.21	2.40	0.89
Word count	20.00	7.04	21.17	6.85	30.50	3.08
Surface elements						
Number of elements	7.80	1.48	7.67	1.03	5.50	1.52
Number of elements repeated by other groups	3.40	1.14	6.33	1.21	4.33	1.03
Percentage of elements repeated by other groups	43.74	13.12	82.54	11.18	80.00	11.30
Literal effects						
Number of effects	3.80	0.84	6.00	2.76	3.83	1.83
Number of effects repeated by other groups	2.20	0.45	3.67	1.63	1.83	1.17
Percentage of effects repeated by other groups	61.33	23.64	67.26	29.31	44.44	22.77
Number of key effects	1.60	1.14	2.67	1.63	0.83	0.98
Number of key effects repeated by other groups	0.80	0.84	2.17	1.17	0.50	0.55
Percentage of effects identified by this group as key	43.00	37.01	46.77	29.51	19.44	22.15
Percentage of key effects identified as key by other groups	43.33	43.46	73.33	43.20	33.33	40.82
Abstract effects						
Number of effects	5.00	1.41	5.33	2.07	4.50	1.05
Number of effects repeated by other groups	2.80	0.84	3.67	1.03	1.83	0.98
Percentage of effects repeated by other groups	56.00	59.16	68.75	50.00	40.74	93.74
Number of key effects	2.00	1.87	2.50	1.38	2.67	1.37
Number of key effects repeated by other groups	1.00	1.00	1.83	1.17	1.17	0.98
Percentage of effects identified by this group as key	35.95	26.81	49.35	26.57	56.94	19.62
Percentage of key effects identified as key by other groups	38.00	41.47	62.50	39.35	40.00	38.87
Distilled experience						
Sentence count	1.60	0.89	1.67	1.03	1.67	0.82
Word count	20.40	9.40	21.67	7.66	20.00	8.29
Context of new experience						
Word count	3.40	2.07	1.33	0.52	9.67	7.20
New experience						
Sentence count	4.50	3.70	5.50	2.74	5.00	1.58
Word count	47.80	17.94	65.33	20.20	57.67	26.23

Table F-2 Numeric aggregate data on TAPT artefacts

	Microblog	Picnic	Wiki
Surface elements			
Number of unique elements	29	19	16
Number of unique elements identified by more than one group	7	11	9
Percentage of unique elements identified by more than one group	24	58	56
Literal effects			
Number of unique effects	11	22	18
Number of unique effects to be identified by more than one group	4	8	5
Percentage of unique effects identified by more than one group	36	36	28
Number of unique key effects	6	10	5
Number of unique key effects to be identified by more than one group	2	4	0
Percentage of unique key effects identified by more than one group	33	40	0
Abstract effects			
Number of unique effects	18	17	20
Number of unique effects to be identified by more than one group	6	7	4
Percentage of unique effects identified by more than one group	33	41	20
Number of unique key effects	8	8	14
Number of unique key effects to be identified by more than one group	2	3	3
Percentage of unique key effects identified by more than one group	25	38	21

Table F-3 Numeric aggregate data on unique elements within TAPT artefacts

	Microblog		**Picnic**		**Wiki**	
	Average	SD	Average	SD	Average	SD
Word count	20.00	7.04	21.17	6.85	30.50	3.08

Table F-4 Word counts of starting experience descriptions

	Microblog		**Picnic**		**Wiki**	
	Average	SD	Average	SD	SD	Average
Number of elements	7.80	1.48	7.67	1.03	5.50	1.52
Percentage of elements repeated by other groups	43.74	13.12	82.54	11.18	80.00	11.30

Table F-5 Per group numeric data on surface elements

	Microblog		**Picnic**		**Wiki**	
	Average	SD	Average	SD	SD	Average
Number of elements	7.80	1.48	7.67	1.03	5.50	1.52
Percentage of elements repeated by other groups	43.74	13.12	82.54	11.18	80.00	11.30

Table F-6 Per group numeric data on surface elements

	Microblog	Picnic	Wiki
Number of unique elements	29	19	16
Percentage of unique elements identified by more than one group	24	58	56

Table F-7 Aggregate numeric data on surface elements

	Microblog		Picnic		Wiki	
	Average	SD	Average	SD	Average	SD
Number of effects	3.80	0.84	6.00	2.76	3.83	1.83
Percentage of effects identified by other groups	61.33	23.64	67.26	29.31	44.44	22.77
Number of key effects	1.60	1.14	2.67	1.63	0.83	0.98
Percentage of effects identified by this group as key	43.00	37.01	46.77	29.51	19.44	22.15
Percentage of key effects identified as key by other groups	43.33	43.46	73.33	43.20	33.33	40.82

Table F-8 Per group numeric data on literal effects

	Microblog	Picnic	Wiki
Number of unique effects	11	22	18
Percentage of unique effects identified by more than one group	36	36	28
Number of unique key effects	6	10	5
Percentage of unique key effects identified by more than one group	33	40	0

Table F-9 Aggregate numeric data on literal effects

	Microblog		Picnic		Wiki	
	Average	SD	Average	SD	Average	SD
Number of effects	5.00	1.41	5.33	2.07	4.50	1.05
Percentage of effects repeated by other groups	56.00	59.16	68.75	50.00	40.74	93.74
Number of key effects	2.00	1.87	2.50	1.38	2.67	1.37
Percentage of effects identified by this group as key	35.95	26.81	49.35	26.57	56.94	19.62
Percentage of key effects identified as key by other groups	38.00	41.47	62.50	39.35	40.00	38.87

Table F-10 Per group numeric data on abstract effects

	Microblog	Picnic	Wiki
Number of unique effects	18	17	20
Percentage of unique effects identified by more than one group	33	41	20
Number of unique key effects	8	8	14
Percentage of unique key effects identified by more than one group	25	38	21

Table F-11 Aggregate numeric data on abstract effects

	Microblog		Picnic		Wiki	
	Average	SD	Average	SD	Average	SD
Word count	20.40	9.4	21.67	7.66	20.00	8.29

Table F-12 Word count of distilled experience descriptions

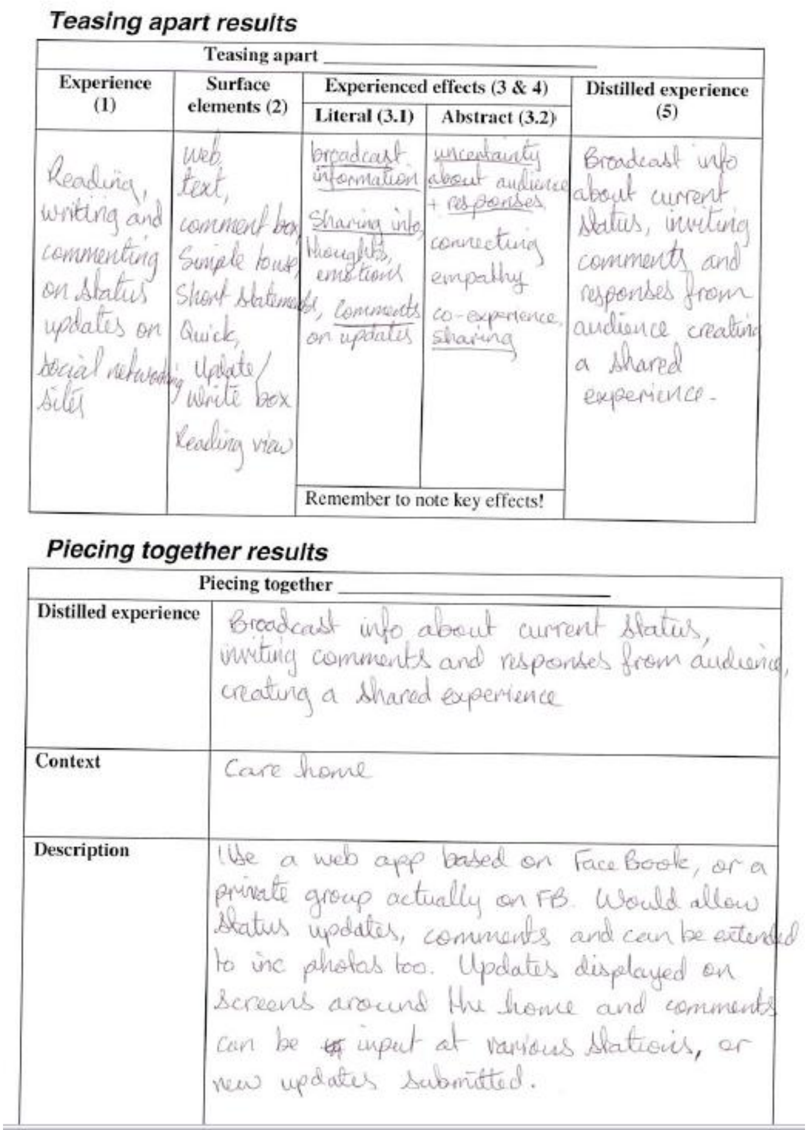
	Microblog		Picnic		Wiki	
	Average	SD	Average	SD	Average	SD
Word count	3.40	2.07	1.33	0.52	9.67	7.20

Table F-13 Word count of descriptions of the context of the new experience

	Microblog		Picnic		Wiki	
	Average	SD	Average	SD	Average	SD
Word count	47.80	17.94	65.33	20.20	57.67	26.23

Table F-14 Aggregate data on the new experience descriptions

Appendix G: Artefacts from the Comparative Evaluation



Teasing apart results

Teasing apart				
Experience (1)	Surface elements (2)	Experienced effects (3 & 4) Literal (3.1)	Abstract (3.2)	Distilled experience (5)
Reading, writing and commenting on status updates on social networking sites	web, text, comment box, Simple touch, Short statements, Quick, Update/Write box, Reading view	broadcast information; Sharing info, thoughts, emotions, comments on updates	uncertainty about audience + responses; connecting, empathy, co-experience, sharing	Broadcast info about current status, inviting comments and responses from audience creating a shared experience.
		Remember to note key effects!		

Piecing together results

Piecing together	
Distilled experience	Broadcast info about current status, inviting comments and responses from audience, creating a shared experience
Context	Care home
Description	Use a web app based on FaceBook, or a private group actually on FB. Would allow status updates, comments and can be extended to inc photos too. Updates displayed on screens around the home and comments can be ~~ip~~ input at various stations, or new updates submitted.

Figure G-1 Sample TAPT artefact: Group K, microblogs

Teasing apart results

Teasing apart Picnic				
Experience (1)	**Surface elements (2)**	**Experienced effects (3 & 4)**		**Distilled experience (5)**
		Literal (3.1)	**Abstract (3.2)**	
A group of people meeting out of doors for a shared meal. During picnic people will converse and will bring food with them to share.	- Food - Picnic rug - Picnic basket - Drink - Sunny - Group of Friends. - Outdoors.	- Food shared around. - Rubbish created. - Conversation	- Companionship - Enjoyment. - Relaxation - Satisfaction	A relaxing and enjoyable shared experience in relaxing surroundings when friends meet to share food and ~~share~~ engage in conversation.
		Remember to note key effects!		

Piecing together results

Piecing together Picnic.	
Distilled experience	A relaxing & enjoyable shared experience in relaxing surroundings when friends meet to share food & converse.
Context	Website
Description	A website that allows friends to 'meet' in a 2D or 3D virtual space. Environment could be customised by host/users eg. beach, field etc. Real time discussion. Ideally would be integrated with other social ~~ne~~ networks eg. FaceBook to allow inviting and creation/scheduling of the event. Product can be selected from company catalogue eg. picnic rug and ordered if desired. Birdsong/background noises played to participant alongside speech.

Figure G-2 Sample TAPT artefact: Group L, picnics

Microblog artefacts generated with TAPT were:

1. Individual handheld devices allow updates via voice input or visual templates; other residents can access and respond to these. (Group B).
2. Interactive TV, menu-based interface to enable phone calls, viewing of news/images, discussions (input method unclear). (Group I).

3. Facebook-based web interface allows status updates, comments and photos: updates displayed on screens around the home, input at ‘various stations’. (Group K).
4. Tablet-based interactions with a ‘very simple UI’: a few large, contextual buttons (e.g. ‘post status’ ‘read’ ‘comment’). Statuses can use templates, text input via on-screen keyboard. (Group N).
5. Input via touchscreen, voice or staff typing it up. Output via common room projector, voice, door-based interface or ‘in room – read and respond from bed’ (unclear). (Group P).

Microblog artefacts generated with Unstructured Discussion were:

1. Communal noticeboard, updated by pinning up index cards (staff remove older material). (Group F).
2. Public touchscreen tablets with icons to build statuses (via a stylus). Private keyboards and tablets. Updates shown on screens in communal areas and private TVs. Private messaging exists. (Group L).
3. Large interactive whiteboard, entry via a touchscreen e-ink tablet (input via handwriting or speech recognition). Viewing on tablets, whiteboard or via audio. (Group Q).
4. Voice-based system with three buttons (‘record message’, ‘hear new statuses’, ‘comment on status’). Built as a hardware device, or integrated into a TV. (Group T).

Microblog artefacts generated with Scenarios and Personas were:

1. Staff-assisted input of text and images, handheld device-enabled input and output, TV viewing of materials. Family access to system. (Group C).
2. Phone-based input and output: four buttons (‘record’, ‘recent messages’, ‘send’ and ‘other messages by this person’) (Group E).
3. Input via keyboard or voice recognition, output via audio or TV screens. (Group H).
4. TV-based input and output: icons to build statuses based on selecting people, events, emotions, animations. Events can be built, as in Facebook. (Group O).
5. Surface displays to write statuses via stylus, finger, keyboard or contextual buttons. Carer access to updates. (Group S).

Picnic artefacts generated with TAPT were:

1. A website to facilitate arranging time and choosing food; delivery of food and drink as ordered by participants; real-time conferencing on website as people eat; ability to comment on picnic afterwards. (Group C).
2. A webpage to let people arrange the event and choose what food to ‘bring’; on the day delivery of food (a selection from the whole list); real-time participation in a Second Life-style virtual world with VOIP. (Group F).
3. An online shop. Users select scenery, and based upon that some products are displayed. Users can interact socially and play web games (e.g. throwing a ball); they can see one another’s shopping baskets and agree to ‘meet up with’ one another. (Group I).
4. Friends ‘meet’ online in a 2D or 3D space, with customised environment such as a beach or field. Real-time discussion occurs, and users can buy products. May be integrated with Facebook. Background noises like birdsong included. (Group L).

5. Friends connect in real-time to a ‘location’ chosen by the host. They share virtual food and optionally participate in activities/chat. It’s a public arena with sound effects and location-themed activities. (Group O).
6. Webpage with tartan background (this background is the same for picnic participant) and webcam portals for each person present. Random alerts with picnic-related problems (bees, rain). (Group Q).

Picnic artefacts generated with Unstructured Discussion were:

1. Beforehand, participants choose which dishes to bring, and pay. On the day, hamper selections and headsets are delivered. Participants gather in a virtual lobby with voice chat, and walk to their virtual location. Voice recognition tracks who’s eating what and their reactions, which is translated to avatar animations. ‘Kill the ants’ style mini-games. Dusk falls as end draws near. (Group D).
2. Host chooses virtual location, picnickers choose menu and food to share (this is delivered). On the day, they log in. Their avatars walk from a virtual car park to the site. Webcams superimpose faces on avatars. The webcam and computer convert picking up real objects to virtual reality. Users share food by inviting others to open specific packets. Ambient sound, views, effects (wandering cows). (Group J).
3. A site with IM / voice / video chatting and a 3D representation of location (countryside, a raft on the ocean, the moon…). Participants can share food (the site suggests items from grocery store websites) and entertaiUDent (streamlining music, virtual Frisbee or football). Blue screens for outdoor visualisation. Product promotion via virtual representations or branded items to buy. (Group P).
4. E-commerce site providing picnic products. VR goggles/motion tracking equipment for picnicking online. Site sells food and accessories (e.g. blankets), delivered the day before the picnic. Each picnic has a limited set of food (recreating the disappointment of missing the last piece of cake). Social media facilities provided. (Group S).

Picnic artefacts generated with Scenarios and Personas were:

1. A tiered system: basic, free client, you pay for better services (webcams etc). Communication via microphone, webcam or text-to-speech. Invitation-only picnics. Configurable clients (e.g. ignore animations/games). Advertising of products in place. (Group A).
2. Friends input preferences about time, location, menu, activities. Site suggests time based on forecast and guest availability. Integrated map and reminders. May suggest activities based on location/group preferences. Guests pick which items to bring. Feedback post-picnic. Site appearance (e.g. location backdrop) reflects choices. Virtual hamper to suggest products based on preferences. (Group G).
3. Flash website with minimal UI. Web cams and microphones for communication. Virtual picnic land with cute, customisable avatars. Some simple games (e-frisbee), file sharing (music, pictures). Event notifications. (Group M).
4. Online shop selling picnic items. First page has categories, search, hot/discounted products. Multiple pictures per product. Products displayed as if you’re having a picnic, hover over items to see their categories, click to go to each item’s product page. Users can edit what products are shown to see how they look together. The system calculates price and recommends other products. (Group R).
5. Website shows sunshine, plays birdsong/bees buzzing, allows you to chat with logged in friends, play flash Frisbee or a fly swatting game. Virtual sausage rolls link to a page to buy real rolls. Can network with picnic fans to find nearby

picnics, and upload picnic photos. Central to the site is the facility to purchase picnic supplies (presented in a pleasing way, consistent with a real picnic, products laid around the picnic area). (Group U).

Wiki artefacts generated with TAPT were:

1. A moderated interactive wall, with access to the history of pages via terminals, and links between comments and areas of the museum. (Group A).
2. A room where visitors can discuss topics, and make changes to 'the agreed version'. (This description was somewhat unclear.) (Group D).
3. A projected wall with enlarged and smaller text, and access to page history and older pages via computers (old comments disappear with time). Computers used for editing, and units available to record/play audio comments. (Group E).
4. An editable wall interface showing linked text, video and audio data. Guests can move elements spatially around the wall. Could link with another museum for a global shared environment. (Group G).
5. Touchscreen maps and timelines, showing videoed user stories as linked with location and time. Facilities provided to record fresh stories. (Group J).
6. Interactive whiteboard with pen-based interface and the ability to add photos. Entries fade (and are archived to a traditional wiki) with time. (Group M).

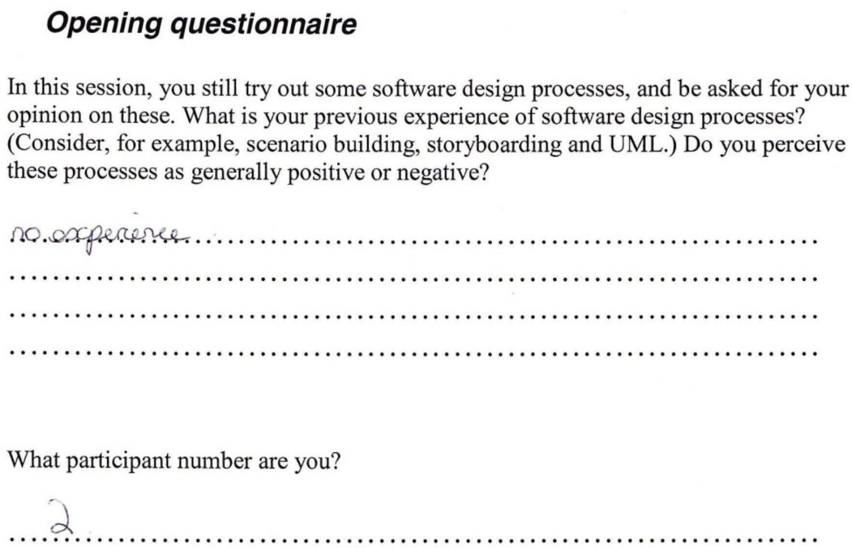
Wiki artefacts generated with Unstructured Discussion were:

1. A display divided into topic areas. Visitors can add/view content via consoles or smartphones, and can voting on articles (popularity or relevance). Voting affects the size of topic areas. (Group B).
2. A virtual visitor book. Visitors can add text / animations / sketches, and comment on / edit entries. Online read access to entries and comments. (Group H).
3. Interactive searchable wall: visitors can add / edit material with a stylus. Experiences can be tagged and annotated. (Group N).
4. Room providing video (not editable), audio (visitors can add audio), virtual wall (visitors may write upon this, comments can overlap), images (not editable). (Group R).
5. Projector display shows latest posts and most replied-to posts in speech bubbles or on bricks. Visitors can post text via mobile phones or terminals: visitors can reply to but cannot edit existing posts. (Group U).

Wiki artefacts generated with Scenarios and Personas were:

1. A two-storey Berlin wall replica. The upper half provides a random data selection, and the ability to select and search for material via touchscreens and Wii type interaction. The lower half enables the addition of data, via either touchscreen, video booth, writing tablets, Wii style input, or digitizing old photos and video. Online view for adding/reviewing data. Links between similar experiences. Timeline available. (Group K).
2. A large display shows a selection of shared experiences. Location-aware mobile devices show personal experiences linked with current artefacts as visitors browse the museum; the devices also allow users to search for and tag experiences, and comment on museum artefacts. (Group T).

Appendix H: Sample Questionnaire Response



Opening questionnaire

In this session, you still try out some software design processes, and be asked for your opinion on these. What is your previous experience of software design processes? (Consider, for example, scenario building, storyboarding and UML.) Do you perceive these processes as generally positive or negative?

no experience..

What participant number are you?

...2...

Figure H-1 A completed opening questionnaire

Questionnaire

What is your participant number?

2

Did you carry out Design Exercise 1, 2 or 3?

1 ☐ 2 ☑ 3 ☐

Did you use Method 1, 2 or 3?

1 ☐ 2 ☐ 3 ☑

Clarity of explanation

The following questions are intended to explore how clear the exercise was.

Were you able to carry out the task? If not, why?

Yes

Were the task and the process explained clearly? If not, how so?

Clearly explained

Did you find any steps of the process difficult? If so, which ones, and how so?

The ~~Tear~~ Teasing Apart phase as the I found it difficult to put into words some of the aspects of a generic wiki environment

Other methods

Do you think there are other processes that you might have used to do this sort of work? Please explain why they are better or worse than the one you applied.

no

Page 1 of 3

Figure H-2 A completed design task questionnaire (page one)

Understanding and replicating experience

The following table is intended to explore how useful the method you used was in helping you understand and re-create experiences. Please write any thoughts you have in the space provided in the second and third columns; the first column gives guidance on what we're looking for.

Ratings: Please provide a rating in answer to each question. Choose from Very Well (VW), Well (W), Okay (OK), Badly (B), Very Badly (VB).

Aspect of interest	Increased understanding	Replication of original experience
Overall…	How well did the method do at improving your understanding of the original experience? How so? It made me analyse what was actually happening Rating: W	How well does your artefact recreate the experience of the original? Does it support experiences analogous to the original? There is a similar sort of experience in both though the tablet original is less specific and less interactive UI W
Emotional aspects. For example, these might include feelings of diligence, happiness or anticipation.	How well did the method do at improving your awareness of emotional effects of the original experience? Why? I made you think about what emotions we experienced when using a wiki that you don't realise you experience. W	How well does your redesigned experience replicate the emotional aspects of the original? Would users experience similar emotional reaction as with the original? I believe the emotional response would be greater as it will evoke emotional responses as it is personal memories as well as it being an actual wall. W VW
The social context and social effects of the experience. For example, people's social expectations beforehand and social fulfilment afterwards.	How well did the method improve your understanding of the social context of the original experience? Why? I had to analyse the social aspects of a wiki. W	How well does your design replicate the social context of the initial experience? In what way? I believe it is better as you can see who wrote what potentially and it is a social experience better than the original W
User perceptions, expectations and reactions during the sequence of events.	How well did the method improve your understanding of the changing perceptions, expectations and reactions of users before, during and after the experience you analysed? In what way? Made you think about using a wiki differently. W	How well will user perceptions, expectations and reactions around the new experience match those of the original? Why? It is a lot more personal in the new method than in the original therefore it doesn't match it well. I believe there's a lot more perceptions, expectations and reactions. O

Page 2 of 3

Figure H-3 A completed design task questionnaire (page two)

More questions on usefulness

How well did the method scaffold a discussion about the problem?

Very well ☐ Well ☑ Ok ☐ Badly ☐ Very badly ☐

Why? users had different opinions on what a wiki was and did not how users reacted to a wiki

Some of the approaches in this study were formal (scenarios and TAPT), and some less so. What was the effect of this structure or lack of structure on the way you worked?

more collaborative rather than individual work.

How well did the method support you in creating imaginative and novel designs?

Very well ☐ Well ☑ Ok ☐ Badly ☐ Very badly ☐

Why? the TA allowed users to gain insight into the key aspects so into the PT users could put a better idea with the same core values and items.

How well are the artefacts you generated (written text, tables, concept maps, scenarios) suited to use in the workplace? (I.e., for documentation or communication.)

Very well ☐ Well ☐ Ok ☑ Badly ☐ Very badly ☐

How so?

not very accessible as entered via query well but could be viewed through a terminal, could be used in group discussion.

How well did you replicate the surface design elements of the original experience? (These are tangible design elements, such as textboxes and graphics in a webpage or shiny wrapping paper and a paper hat in a Christmas cracker.)

Very well ☐ Well ☐ Ok ☐ Badly ☑ Very badly ☐

Why?

we redesigned the look and feel to the original totally which worked for the new one but not for the original

Considering the process as a whole, do you have any other comments about its usefulness or value?

Very usefull, allowed users to understand the original before creating the new version.

Page 3 of 3

Figure H-4 A completed design task questionnaire (page three)

Closing questionnaire

Considering the methods you have applied today, which did you find easier to use? (Consider how straightforward the methods were, and how long they took to apply.) Why?

T.A.P.T. made me analyse the original code fully understand it before I tried to make something new or improve it

Which did you feel produced more useful (powerful, practical, relevant) results? Why?

TAPT. Because you understood what the core elements of the original were before creating the first prototype it enlarged and allowing more powerful and practical methods to be created

Which did you feel was more effective at replicating the experience you were considering? Why?

TAPT. All elements were thought out and analysed before theories are new created

What participant number are you?

2

Finally, a big **thank you** for your time and effort. It is very much appreciated!

Figure H-5 A completed closing questionnaire

Appendix I: Comparative Evaluation Ratings of Methods

Table I-1 shows participant ratings of methods in the comparative evaluation. A key to the questions is below.

Question	**Method**														
	TAPT					**Scenarios**					**Unstructured Discussion**				
	Very well	Well	Ok	Badly	Very badly	Very well	Well	Ok	Badly	Very badly	Very well	Well	Ok	Badly	Very badly
1	10	14	7	1	0	4	8	7	5	0	5	7	10	4	1
2	8	14	6	3	1	2	9	7	5	1	3	7	7	5	5
3	7	11	8	4	1	2	9	8	4	0	4	4	12	4	2
4	2	12	8	7	1	2	8	8	5	1	0	2	12	9	4
5	7	9	11	4	0	2	13	7	2	0	2	17	6	1	1
6	1	19	7	4	0	1	14	6	2	1	3	9	7	7	0
7	4	13	9	3	0	4	8	8	3	0	5	12	6	3	1
8	2	12	14	2	0	1	11	7	4	0	1	6	12	6	1
9	12	18	0	2	0	8	10	5	1	0	4	6	6	6	5
10	8	10	11	2	0	6	9	7	2	0	11	11	4	0	1
11	4	12	12	4	0	2	11	10	1	0	2	1	12	8	3
12	5	11	7	5	1	1	10	8	2	2	6	8	4	1	7

Table I-1 Comparative evaluation ratings

Questions:

1. How well did the method do at improving your understanding of the original experience? How so?
2. How well did the method do at improving your awareness of emotional effects of the original experience? Why?
3. How well did the method improve your understanding of the social context of the original experience? Why?
4. How well did the method improve your understanding of the changing perceptions, expectations and reactions of users before, during and after the experience you analysed? In what way?
5. How well does your artefact recreate the experience of the original? Does it support experiences analogous to the original?
6. How well does your redesigned experience replicate the emotional aspects of the original? Would users experience similar emotional reaction as with the original?
7. How well does your design replicate the social context of the initial experience? In what way?
8. How well will user perceptions, expectations and reactions around the new experience match those of the original? Why?
9. How well did the method scaffold a discussion about the problem?
10. How well did the method support you in creating imaginative and novel designs?
11. How well are the artefacts you generated (written text, tables, concept maps, scenarios) suited to use in the workplace? (I.e., for documentation or communication.)
12. How well did you replicate the surface design elements of the original experience? (These are tangible design elements, such as textboxes and graphics in a webpage or shiny wrapping paper and a paper hat in a Christmas cracker.)

Appendix J: Expert Review Materials

Questionnaire

On the design task

1) What would you say are the key challenges in this task? [NB note these down]
2) What kind of approach might you use in response to this task?

On each individual artefact

1) Can you characterise this artefact? For example, is it surprising, insightful, misleading, meaningful, good?
2) Does the designer exhibit any unusual understanding or perception? Were they innovative? Why?
3) Do you think the designer has responded to the key challenges in the task?
4) Does the artefact translate the experience? (And is that translation of the deeper, underlying experience – for example, of emotional or social aspects of the experience – or is it a translation of more superficial, design elements?)
5) Do the designers seem to have carried out significant analysis in order to produce this artefact? Do you think they were aware of any assumptions they may have made?
6) Is the artefact inclusive or accessible?
7) Have you any other comments on this artefact?

On all individual artefacts

1) Did you notice any patterns or themes running through the artefacts you have seen today? What? Why do you think they are present?
2) Were you surprised by the homogeneity (or lack thereof) across artefacts?
3) Do either of the artefacts you reviewed stand out to you as especially inspired or dull? Why?
4) Have you any other comments about the artefacts you have seen?

On each individual TAPT artefact

1) Is your perception of the design artefact different now you've seen the analysis which was conducted towards building that description? Why?
2) Given the analysis shown here, would you have come up with a significantly different design artefact? Why?
3) The first table listed lists surface elements, experienced effects and the distilled experience. Were you surprised by any of that content? Were any contradictions present?
4) Were you surprised by what was, or was not, chosen as key?
5) Do you have any other comments about this artefact?

On both TAPT artefacts

1) Did you notice any patterns of themes running through the TAPT artefacts you have seen today? What? Why do you think they are present?
2) Were you surprised by how much or by how little the different analyses overlapped in the choice of elements, effects and key effects?
3) Was there much variety in the emphasis of the analyses, and what do you think that means? For example, maybe one had more surface elements and the other more abstract effects, or maybe one chose only one or two effects as key while the other chose many.
4) Do any of the analyses you reviewed stand out to you as especially inspired or dull? Why?
5) Have you any other comments about the TAPT artefacts you have seen?

Participant Information Sheet

Study Title: Evaluation of the TAPT process
Researcher: Clare Hooper
Ethics number: ES/10/06/001

Please read this information carefully before deciding to take part in this research.

If you are happy to participate you will be asked to sign a consent form.

What is the research about?

This study is designed to evaluate several different software engineering design methods. In a previous experiment, participants carried out design tasks using one of three different approaches: this experiment involves evaluating the resultant design artefacts.

Why have I been chosen?

You have been approached to participate in this study because you are an expert in one of the domain areas of the design artefacts (these are: wiki design; microblogging design; social interactions). As such, you are well-placed to comment on the efficacy of the artefacts.

What will happen to me if I take part?

You will be shown the design task given to participants of the previous experiment. You will be shown the resultant design artefacts (which will be no more than 15 in number, and each consist of several hundred words of text). You will be asked for your opinion of the artefacts. The interview will be subject to audio recording, and will take no more than one hour.

Are there any benefits in my taking part?

By taking part, you have the opportunity to gain insight into laypeople's responses to simple design tasks. You may view the study as an opportunity to exercise your evaluative skills.

Are there any risks involved?

No.

Will my participation be confidential?

Yes. Audio recordings will be edited to assure anonymity: if, for example, a person's name is mentioned, the stored data will be edited to remove this detail.

What happens if I change my mind?

You have the right to withdraw at any time without your legal rights being affected.

What happens if something goes wrong?

If you have a concern or complaint, please contact the ECS School Office on school@ecs.soton.ac.uk or school@ecs.soton.ac.uk02380 592909.

Where can I get more information?

Feel free to get in touch with Clare Hooper (clare@ecs.soton.ac.uk, 02380 597208) or her supervisor, David Millard (dem@ecs.soton.ac.uk, 023 8059 5567).

Who is sponsoring this research?

The research sponsor is the School of Electronics & Computer Science at the University of Southampton.

Consent Form

Study title: Expert review of design artefacts

Researcher name: Clare Hooper

Ethics reference: ES/10/06/001

Please initial the box(es) if you agree with the statement(s):

I have read and understood the information sheet (June 2010 / Version 1.0) and have had the opportunity to ask questions about the study. ☐

I agree to take part in this research project and agree for the resultant data to be used for the purpose of this study. ☐

I understand my participation is voluntary and I may withdraw at any time without my legal rights being affected. ☐

I agree to read a set of brief design artefacts, and answer some verbal questions about my opinion on these. ☐

I agree to the recording of above interview: I understand that no data which can identify me will be stored. ☐

Name of participant (print name)……………………………………………………

Signature of participant…………………………………………………………..

Name of Researcher (print name) ……………………………………………………

Signature of Researcher…………………………………………………………..

Date……………………………………………………………………………

Appendix K: Sample Expert Review Transcript

The below transcript conveys an early stage of the interview with Participant A of the expert review (who examined microblog artefacts). "…" refers to a pause in speech.

Time	Speaker	Discussion
01:09	Researcher	I'm going to anonymise the results … it would be really helpful if we can agree a title … something to reflect your expertise … so something like second year PhD student
	Participant	yep
	Researcher	studying blogs?
	Participant	yep
	Researcher	awesome … so onto the task … two questions about the task itself … what would you say would be the key challenges in doing this task?
	Participant	for me the first thing that pops up is usability … because you've got a really specific kind of audience right there so how they can use it … and possibly the conventions … so the conventions are going to be different as well
	Researcher	how do you mean?
02:00	Participant	'cos basically all the arty stuff … all the conventions that happen in Twitter … it's basically involving from the audience themselves … even Twitter themselves admitted that it's not them that said "ok, if you want to copy and then attribute things you do a retweet" that's sort of like the whole world who does that … so if the audience is different that means they all have their own conventions of how they do things … how they want to use it as well
	Researcher	so it's supporting them when they don't necessarily know…?
	Participant	yep yep
	Researcher	ok … conventions cool … we'll come back to that … and one other thing about the task is … if you were doing this, what sort of method would you use like how would you go about it?
	Participant	how do you mean method? like the … developing or…?
	Researcher	er the designing … if you were coming up with the system

Appendix L: NVivo Screenshots

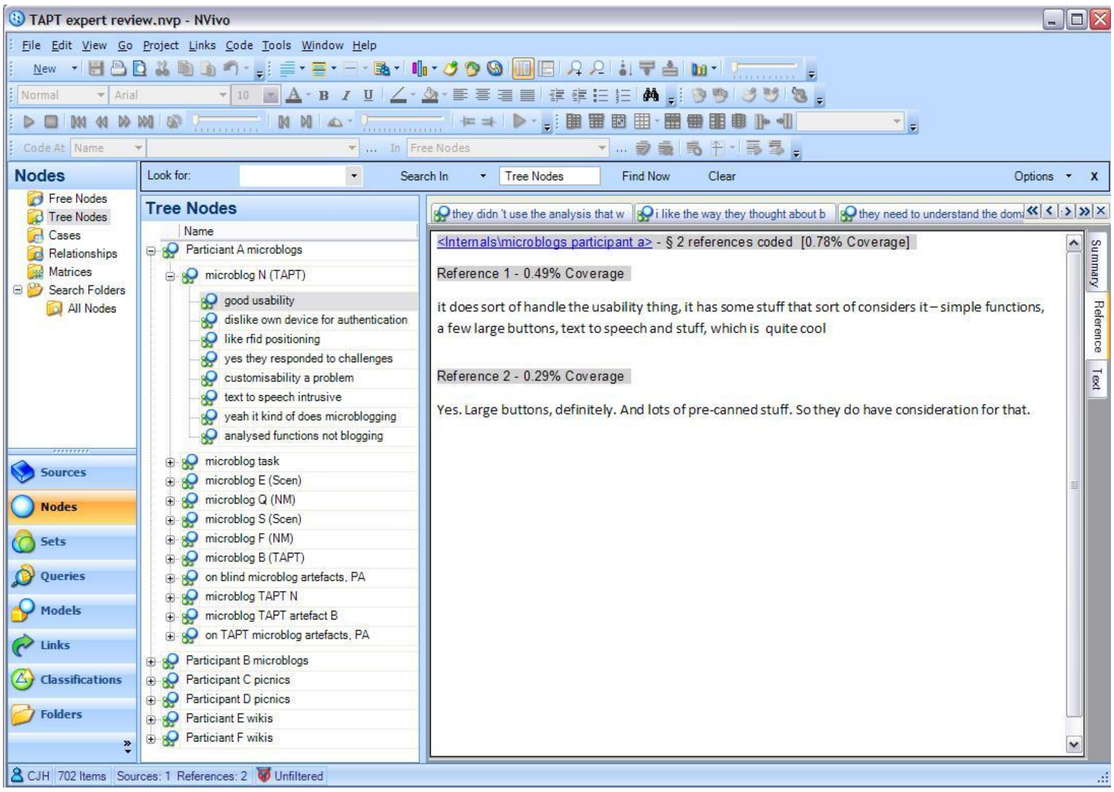


Figure L-1 Nodes in NVivo

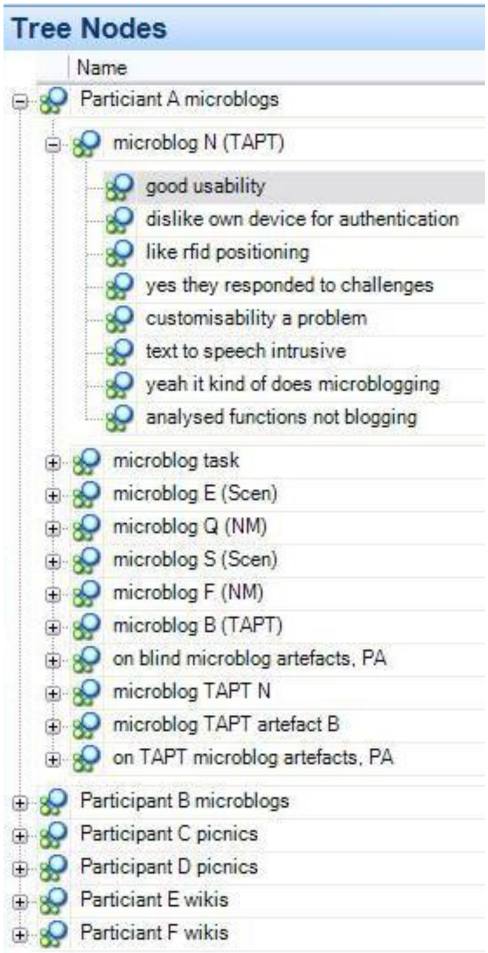


Figure L-2 Close up of nodes in NVivo

Appendix M: Numeric Expert Review Data

	Average	Microblog	Picnic	Wiki
TAPT	69	56	76	73
Scenarios	105	134	85	97
Unstructured Discussion	129	116	135	136

Table M-1 Word counts of normalised artefacts

Expert	Artefact					
	First	Second	Third	Fourth	Fifth	Sixth
A	0.8	0.7	1.0	0.6	0.9	0.8
B	0.2	0.5	0.0	0.4	0.2	0.6
C	0.0	0.0	0.2	0.0	0.0	0.1
D	0.3	0.8	0.5	0.9	0.3	0.9
E	0.1	1.0	0.2	0.9	0.8	0.0
F	0.1	0.3	0.5	0.0	0.0	0.4
Average	0.2	0.6	0.4	0.5	0.4	0.5

Table M-2 Artefact scores over time

Method	Task	Artefact	Score	Expert	Artefact provenance
TAPT	Microblog	B	0.83	A	IBM
		N	0.75		IBM
		K	0.20	B	IBM
		N	0.40		IBM
	Picnic	C	0.17	C	IBM
		L	0.1		IBM
		F	0.3	D	IBM
		L	0.3		IBM
	Wiki	A	0.88	E	IBM
		J	0.2		IBM
		A	0	F	IBM
		G	0.4		IBM
Scenarios	Microblogs	E	0.70	A	IBM
		S	0.60		University
		E	0.00	B	IBM
		O	0.20		IBM
	Picnic	M	0	C	IBM
		R	0		University
		A	0.5	D	IBM
		U	0.9		University
	Wikis	K	0.1	E	IBM
		T	0		University
		K	0.5	F	IBM
		T	0		University
Unstructured Discussion	Microblog	F	0.90	A	IBM
		Q	1.00		University
		F	0.50	B	IBM
		L	0.60		IBM
	Picnic	P	0.00	C	IBM
		S	0.00		University
		J	0.8	D	IBM
		S	0.9		University
	Wikis	N	0.80	E	IBM
		U	1.00		University
		B	0.10	F	IBM
		R	0.30		University

Table M-3 Artefact scores

Microblogs	TAPT	Scenarios	Unstructured Discussion	Average
Innovation	0.3	0.5	0.8	0.5
Challenges	0.7	0.4	0.8	0.6
Translation	0.4	0.3	1.0	0.5
Analysis	0.3	0.3	0.6	0.4
Accessible	0.8	0.5	0.6	0.6
Average	0.5	0.4	0.8	0.6

Table M-4 Microblog scores (by category)

Picnics	TAPT	Scenarios	UD	Average
Innovation	0.2	0.5	0.7	0.4
Challenges	0.3	0.4	0.4	0.3
Translation	0.4	0.4	0.5	0.4
Analysis	0.0	0.4	0.4	0.3
Accessible	0.3	0.3	0.4	0.3
Average	0.2	0.4	0.5	0.3

Table M-5 Picnic scores (by category)

Wikis	TAPT	Scenarios	UD	Average
Innovation	0.4	0.1	0.6	0.4
Challenges	0.4	0.3	0.6	0.4
Translation	0.4	0.3	0.5	0.4
Analysis	0.4	0.0	0.5	0.3
Accessible	0.3	0.1	0.5	0.3
Average	0.4	0.2	0.6	0.4

Table M-6 Wiki scores (by category)

	TAPT	Scenarios	UD	Average
Innovation	0.3	0.4	0.7	0.4
Challenges	0.4	0.3	0.6	0.4
Translation	0.4	0.3	0.7	0.4
Analysis	0.2	0.2	0.5	0.3
Accessible	0.4	0.3	0.5	0.4

Table M-7 Artefact scores

Appendix N: Expert Review Quotations

Quotations on the disconnect between Teasing Apart and Piecing Together:

- Participant A, microblog-B ("[the analysis] basically says a lot about how people use it and what the feel about it but the description is just about what the people can do")
- Participant A, microblog-N ("they underline sense of community but that wasn't really explicit in the [design]. what I understood from the [design] is it's more about how you use it, how to post your status, use of pre-canned sentences. But […] it doesn't really explain how they did the sense of community bit, about how they addressed it.")
- Participant B, microblog-K ("[the design] doesn't reflect all the analysis that they did" and "'connecting', 'empathy', I can see a link there with emotions but it doesn't make its way into the description of the project and it's not explored sufficiently")
- Participant C, picnic-C ("it's odd, there's some disjunction between [the analysis and the design] […]there's some odd gap between the two" and "they're looking at the superficial objective aspects […]but they're not finding it easy to unpack what a shared social experience is")
- Participant C, picnic-L ("I'd have liked them to articulate [the abstract effects] more, these things more in the design artefact, rather than hiding them.")
- Participant D, picnic-F ("They considered things like the emotional aspects, the abstract aspects, more than I thought")
- Participant D, picnic-L ("again they've already written some of that down anyway but they've not taken it further I guess". Participant D was asked, "so another one where the analysis isn't fully reflected in the design?" and she replied "Yeah … it's a bit implicit in there. They've not always addressed it specifically.")
- Participant E, wiki-A ("there's a 'clear way of editing information' [in the analysis]: actually, they don't mention how they're going to get people to edit in the design.")

- Participant E, wiki-J ("I don't see how [the design] uses what they identified , 'Unease' […] 'Satisfaction' […] [the design] doesn't address any of those issues." He went on to say "I can't believe they came up with the description after that analysis. I refuse to believe that that came from this.")
- Participant F, wiki-A ("we want people to have these experiences when they're using it, I'm not sure where they've thought that what they've offered will give those experienced effects. […] where are the experienced effects to make people want to do it? Except for the text being graffiti-like. That's the only thing really.")
- Participant F, wiki-G ("I think [that in the design] they've lost some of the good stuff that they put in [the analysis]" and, of abstract effects "they haven't used the words like 'satisfaction' or some of the other things, like finding out surprising things. So again, this is very much wiki-standpoint: the fact that a wiki is collaborative, that you can exchange knowledge, the community aspect they've understood. But the deeper satisfaction of why you might want to do it… not so much.")

Appendix O: Case Study Materials

Participant Information Sheet

Study Title: Investigating the use of the TAPT process
Researcher: Clare Hooper Owens
Ethics number: ES/10/03/010

Please read this information carefully before deciding to take part in this research. If you are happy to participate you will be asked to sign a consent form.

What is the research about?
The purpose of this study is to evaluate the software engineering method called TAPT, or Teasing Apart, Piecing Together. This is to be achieved by asking participants about their experiences applying TAPT.

Why have I been chosen?
You have been approached to participate in this study because you are a software engineering professional who has some experience with TAPT.

What will happen to me if I take part?
Participants will be asked open-ended verbal questions about how they have used TAPT, and whether it was useful. The interview may be subject to audio recording. Information will be kept private and transcripts will be anonymised.

The interview should take no more than 30 minutes. You may be asked to participate in follow-up interviews in the future, but there is no obligation to do so.

Are there any benefits in my taking part?
You will have the opportunity to reflect upon your use of a new software engineering design technique and to contribute to current knowledge.

Are there any risks involved?
No.

Will my participation be confidential?
Yes. Audio recordings will be edited to assure anonymity: if, for example, a person's name is mentioned, the stored data will be edited to remove this detail. Transcripts of the recordings will reflect this.

What happens if I change my mind?
You have the right to withdraw at any time without your legal rights being affected.

What happens if something goes wrong?

If you have a concern or complaint, please contact the ECS School Office on school@ecs.soton.ac.uk or school@ecs.soton.ac.uk02380 592909.

Where can I get more information?

Feel free to get in touch with Clare or her supervisor, David Millard:

Clare: clare@ecs.soton.ac.uk, 01962 816863 (x246863)

David: dem@ecs.soton.ac.uk, 023 8059 5567

[April 2010] [Version 1.0]

Consent Form

Study title: Investigating the use of the TAPT process

Researcher name: Clare Hooper Owens

Study reference: TAPT case study

Ethics reference: ES/10/03/010

Please initial the box(es) if you agree with the statement(s):

I have read and understood the information sheet (April 2010 / Version 1) and have had the opportunity to ask questions about the study ☐

I agree to take part in an interview for this research project and agree for my data to be used for the purpose of this study ☐

I agree for audio recording of the interview. ☐

I understand my participation is voluntary and I may withdraw at any time without my legal rights being affected ☐

Name of participant (print name)………………………………………………

Signature of participant……………………………………………………..

Name of Researcher (print name) ………………………………………………

Signature of Researcher……………………………………………………..

Date……………………………………………………………………

[April 2010] [Version 1]

Opening questions

Blapr

1) When did you use TAPT?
2) What experience did you tease apart, and into what context did you piece together the experience? Why did you choose that experience and context?
3) Was this in a group? If so, who else was present?
4) How long did it take?
5) Can you tell me a little bit about how it worked? What happened, and why?
6) Did your use of TAPT sit naturally within the design process? Why?
7) Do you think the analysis you carried out during Teasing Apart is useful to your project? Why?
8) Do you think the scenario(s) generated in Piecing Together is useful to your project? Why?
9) At this stage of your project, what (if any) impact do you think your use of TAPT has had on your work? Why?
10) What (if any) impact do you think your use of TAPT will have on your ongoing work? Why?
11) Have you got any final comments?

(Questions 1 – 5 on practical aspects, 6 – 10 on perceptions of TAPT's impact.)

Spoken Web (pre-study)

1) What drove your decisions on how to run this study? For example, how many (and what type) of people to involve, the timing of the sessions.
2) What made you choose TAPT as a tool?
3) Are there any other methods you'd consider choosing for this task? If so, what are they and will you use them as well as TAPT?
4) Are you expecting to get results from TAPT that other methods wouldn't get you? Why?
5) What do you hope to achieve from this study? What are your goals? Why?
6) What impact do you think this study will have on your work?
7) Do you think this study might change your perceptions or understanding of social interactions on the web?
8) Do you think this study might influence your design decisions when building the social spoken web site?
9) Have you any final comments?

(Question 1 on pre-study decisions, 2 – 4 on the choice of TAPT as a tool, 5 – 8 on hopes for the study.)

Spoken Web (immediately post-study)

1) When did you use TAPT, and how long did it take?
2) What experience was teased apart, and into what context was it pieced together? Why did you choose that experience and context?
3) You chose the composition of the groups of people who carried out this work: did the groups work as you expected? Would you make different decisions were you to run the study again?
4) Did the study run as you expected?
5) Can you tell me a little bit about how it worked? What happened, and why? What were the results from this study?
6) Did your use of TAPT sit naturally within the design process? Why?
7) Do you think the results from using TAPT are useful to your work? Why?
8) What (if any) impact do you think your use of TAPT will have on your ongoing work? Why?
9) Do you have new insight into the working of social websites?
10) Before the study, you identified a hope that the work would help you understand more about what people get out of social networking and also get information to drive your voice-based designs. Have you made progress on those two aims?
11) How would you say TAPT compared with other processes you've used to get understanding in this area? What do you think TAPT revealed that other processes would not, and why do you think TAPT revealed these things?
12) Have you any final comments?

(Questions 1 – 5 on practical aspects, 6 – 11 on perceptions of TAPT's impact.)

Gaming

1. What changes have you made to the TAPT instructions, and why?
2. What changes have you made to the blank TAPT form, and why?
3. What drove your choice of games and genres for our participants to tease apart? Why?
4. What made you choose TAPT as a tool?
5. Are there any other methods you'd consider choosing for this task? If so, what are they and will you use them as well as TAPT?
6. Are you expecting to get results from TAPT that other methods wouldn't get you?
7. What do you hope to achieve from this study? What are your goals? Why?
8. What impact do you think this study will have on your work? Do you think it might change your perceptions or understanding of gaming (or people's views of gaming) in some way?
9. Have you any final comments?

(Questions 1 – 3 on pre-study decisions, 4 – 6 on the choice of TAPT as a tool, 7 – 8 on hopes for the study.)

Location-based tools (pre-study)

1. What drove your decision to investigate location-based services such as Gowalla and geocaching?
2. What made you choose TAPT as a tool?
3. Are there any other methods you'd consider choosing for this task? If so, what are they and will you use them as well as TAPT?
4. Are you expecting to get results from TAPT that other methods might not get you?
5. What do you hope to achieve from this study? What are your goals? Why?
6. What impact do you think this study will have on your work? Do you think it might change your perceptions or understanding of location-based services in some way?
7. Have you any final comments?

(Question 1 on motivation, 2 – 4 on the choice of TAPT as a tool, 5 – 6 on hopes for the study.)

Location-based tools (immediately post-study)

1. What were the results from this study?
2. Have you new insight into the POV of the participants, or into how location-based services work?
3. Did the study run as you expected?
4. Did people tease apart the experiences in the way you expected?
5. We jointly made some decisions about the groups of participants: I provided input about how many might work, and you recruited the participants. Did the groups work as you expected? Would you make different decisions were you to run the study again?
   Have you any final comments?

(Questions 1 – 2 on initial insights; 3 – 4 on if expectations were met; 5 about how the study ran.)

Closing Questions

Blapr

1. Did your HBGO project run as you expected?
2. You first used TAPT in April to try and understand systems similar to your own – you looked at Google maps, signposts and a tour guide. Now, seven months on, do you think those analyses drove your resultant designs?
3. How would you say using TAPT compared with other processes you might use to drive designs or understand users and their experiences? What other processes are you aware of?
4. Would you say using TAPT has affected your work outcomes? In what way?
5. Did TAPT help you design things in an original way rather than doing a straight replication of, for example, the functionality of Google maps?
6. Did TAPT help you decide and focus on what struck you as important in this work?
7. Did your use of TAPT sit naturally within the design process? That is, did it do what you needed, when you needed it?
8. Did TAPT meet your expectations as a design tool? Would you use TAPT for design again? Would you change the way in which you used it?
9. You have also used TAPT as part of your evaluation process. Tell me a bit about how you went about that.
10. Did Teasing Apart the Blapr experience help you get new insights into the system?
11. Did comparing the analysis of Blapr to your previous analyses of Google maps, signposts and tour guides help you gain you insights into your system?
12. How would you say TAPT compared with other evaluation processes? What other processes are you aware of?
13. Did you use of TAPT sit naturally within the evaluation process? That is, did it do what you needed, when you needed it?
14. Did TAPT meet your expectations as an evaluative tool? Would you use TAPT for evaluation again? Would you change the way in which you used it?
15. Have you any final comments?

(Questions 2 – 8 on TAPT as a design tool. Questions 9 – 14 on TAPT as an evaluation tool: these were only asked of Participant D, who conducted the evaluation alone.)

Spoken Web

1. Could you briefly describe how you work has gone since we met in August?
2. Did your work go as you expected?
3. In August, we asked a focus group of social network users to tease apart two facets of social networking, status updates and notifications. Before the study, you said you hoped to increase your understanding of what people get out of social networking. After the study, you felt you had a greater understanding of people's social networking experiences, and that that might affect your approach. Did it?
4. You expressed a hope that the study would help you acquire information to drive your voice designs. Now, three months on, do you think you got information to drive those voice designs?
5. On the day, you said that if you got good results from using TAPT in this way, you wouldn't need to do a user survey. Have you since felt the need to conduct such a survey? If so, what data did you require which you didn't get in August?
6. In August, you commented on the value of the abstract effects which participants identified. You said that you intended to look at 'structuring' that abstract experience and trying to model it. Did you conduct work along these lines?
7. Did TAPT meet your expectations?
8. Three months on, how would you say using TAPT compared with other processes you might use to get understanding into people's experience of social networking?
9. Would you say using TAPT has affected your work outcomes? In what way?
10. Did your use of TAPT sit naturally within the design process? That is, did it do what you needed, when you needed it?
11. Would you use TAPT again? Would you change the way in which you used it?
12. Have you any final comments?

(Questions 1 – 2 on results, 3 – 7 on expectations, 8 – 11 on TAPT and other processes)

Gaming

1. Did the study run as you expected?
2. Did people deconstruct games in the way you expected?
3. Would you run the study differently if you were doing it again?
4. What were the results from this study?
5. Have you new insight into the POV of the participants? Would you say the output is of practical use for designing game experiences, as you hoped? (Participant I)
6. Are the analyses useful in their own right? (e.g. by seeing what words people choose to describe things) (Participant I)
7. Did the results help you identify the main experiences or elements involved in games? (Participant H)
8. Do you feel better informed as to what might be appropriate components or elements of educational games? (Participant H)
9. Do you think you have insight into the experiences or elements of different genres as a result of this? (both)
10. Can you use this output towards designing some kind of useful and fun educational games? (both)
11. How would you say using TAPT compared with other processes you've used to get understanding into games and genres of game?
12. Would you say TAPT revealed things that other processes might not? If so, why do you think TAPT revealed these things?
13. Did your use of TAPT sit naturally within the research process? By this, I mean, in the context of conducting a piece of research, did it do what you needed, when you needed it?
14. Have you any final comments?

(Questions 1 – 4 on results; 5 – 10 on expectations (source of question in brackets); 10 – 13 on the method/fit in the research process)

Location-based tools

1. What were the results from this study?
2. Have you new insight into the POV of the participants, or into how location-based services work?
3. Were the analyses produced by participants useful to your work? Why?
4. What (if any) impact do you think your use of TAPT will have on your ongoing work in this area? Why?
5. Did you gain an insight into what Gowalla/Geocaching are to experience on a deeper level? What about insight into /why/ they're fascinating or compelling?
6. Have you any further thoughts about whether you'd run this study differently if repeating it?
7. How would you say using TAPT compared with other processes you've used to understand people's perspectives and experiences? Would you say TAPT revealed things that other processes might not? If so, why do you think TAPT revealed these things?
8. Did your use of TAPT sit naturally within the research process? By this, I mean, in the context of conducting a piece of research, did it do what you needed, when you needed it?
9. Have you any final comments?

(Questions 1 – 4 on results; 5 on expectations; 6 – 8 on the method/fit in the research process)

Appendix P: Sample Case Study Transcript

The below transcript conveys an early stage of the opening interview with case study Participant A (Blapr). "…" refers to a pause in speech.

Time	**Speaker**	**Discussion**
00:30	Researcher	could you tell me when it was … I appreciate it may be more than once … but sort of how long ago you used TAPT?
	Participant	we did it on two Friday afternoon sessions … one of them would've been two weeks ago this coming Friday and the other one was three weeks ago
	Researcher	ok … could you tell me what experience you were piecing … Teasing Apart and the context into which you were piecing it together?
01:00	Participant	in the first session we tried it with Google maps … and then the second one we did a sign post and a tour guide
	Researcher	cool … ok …. and why was it you chose those experiences to tease apart?
	Participant	because we're… we've already got the idea that we wanted to do some kind of routing mapping … um final thing… um we were going with things that we thought would be relevant and different parts of what we wanted to come up with … things like a signpost and maybe a tour guide for extra information things like that would be relevant to our final idea
	Researcher	that makes sense … and what was the context that you pieced those three different bits back together into?
	Participant	um … well with … with the Google maps … it was the first time we'd done it and it was kind of … it didn't work too well … because we were tearing apart [sic] software and putting it together and kind of trying to do it as software again and that didn't work so well … which is why we kind of went with the sign post and the tour guide the more physical things that aren't software at all
	Researcher	yeah because one of you guys mentioned this last week
	Participant	yeah
02:00	Researcher	so … can you explore that a bit more … like why do you think it was that going from software to software was more difficult?

Appendix Q: Case Study Artefacts

Blapr

User experience

1. General employee
◦ A general employee might be a meeting attendee or chair – both have similar requirements.
◦ The meeting chair may have a half-full meeting room and want to know whether or not to expect more people and delay the start of the meeting. To find out, he might use a sametime bot to query the location of a specific, important user or an online map tool to view the location of multiple people at once.
◦ The meeting attendee may have to go from one meeting straight to another without a laptop. To save going back to her desk to find the location of her next meeting and either leave one meeting early or arrive at another late, she can use a mobile device to see where she needs to be on a map.
2. Employee with special access
◦ Some employees may be able to go through restricted areas which are not accessible to everyone. In this case, when routing the employee to a given location, the application should be aware that a quicker route could be through an area not normally considered.
3. Customers
◦ Customers may rely on the system far more than others due to an unfamiliarity with the site. In this case, a customer might need to get from the Executive Briefing Centre to Galileo building or find the nearest accessible toilet. The most appropriate solution would be to use a mobile device to look up a mobile-compatible web page to avoid any cumbersome install/set up times for a single-use occurrence.
4. Disabled
◦ Disabled users could have a very wide range of requirements, but the principle ones would be, as with the employee with special access, the route taken could be very different to allow for use of lifts or ramps and to avoid stairs for those in wheelchairs.
◦ Blind users would benefit from an audible method of receiving information, which would include warnings about potential hazards.

Not all of the above can feasibly be included, but the aim is that the majority of most commonly used functions will be included.

Personas / Scenarios

Persona

- General employee
- Employee with special access
- Customers
- Disabled

Scenario

- Find location of a meeting room
- Applicable to all personas
- Make use of E-ink displays
- Web based map showing all meeting attendees and location of meeting
- Find location of a person/people
- Applicable to all personas
- Web based map showing location
- E-ink displays to a lesser extent
- Mobile devices
- Route to location
- Applicable to all
- 2: can route through restricted areas
- 1,3: restricted travel
- As above, taking lifts/ramps (or the absence thereof) into account
- Audio
- Web based map
- Mobile devices

<table>
<tr><th colspan="5">Teasing apart Google Maps</th></tr>
<tr><th rowspan="2">Experience (1)</th><th rowspan="2">Surface elements (2)</th><th colspan="2">Experienced effects (3 & 4)</th><th rowspan="2">Distilled experience (5)</th></tr>
<tr><th>Literal (3.1)</th><th>Abstract (3.2)</th></tr>
<tr><td rowspan="2">Find a location, view on a map, get directions between 2 or more locations. View route as list of directions or on a map. Can view traffic conditions. Can view walking, car, public transport directions.</td><td rowspan="2">Pictoral, text view of directions, 'basic' view (basic html), zoom function, clear bold colour use (e.g. road, river), option to modify route, opt to alter default location</td><td>broadcast visual information (location + route), showiness of locations, directivity, route text shown in clear steps, lookability of pictures of location</td><td>funness of streetview, suspense, advertise (geo-caching), reminiscence</td><td rowspan="2">a way to convey location information and a route between multiple locations. this information can be consumed in several formats.</td></tr>
<tr><td colspan="2">Remember to note key effects!</td></tr>
</table>

Table Q-1 Blapr: Teasing Apart Google maps

<table>
<tr><th colspan="5">Teasing apart Signpost</th></tr>
<tr><th rowspan="2">Experience (1)</th><th rowspan="2">Surface elements (2)</th><th colspan="2">Experienced effects (3 & 4)</th><th rowspan="2">Distilled experience (5)</th></tr>
<tr><th>Literal (3.1)</th><th>Abstract (3.2)</th></tr>
<tr><td rowspan="2">A notice providing the direction and distance from itself to another location</td><td rowspan="2">Arrow
Bold
<u>Quick to read</u>
Simple shape/colour</td><td><u>Broadcast visual info</u>
<u>Direction</u>
Recursive
Shows distance away from target</td><td>Depressed or excited
<u>Informed</u>
Safety / relief
Confusion
Reminiscent</td><td rowspan="2">A way to broadcast information that informs the user of their <u>relationship</u> between where they are and <u>where they are going</u> leaving them in a wide range of <u>emotional states</u></td></tr>
<tr><td colspan="2">Remember to note key effects!</td></tr>
</table>

Table Q-2 Blapr: Teasing Apart a signpost

<table>
<tr><th colspan="2">Piecing together Signpost</th></tr>
<tr><td>Distilled experience</td><td>A way to broadcast information that informs the user of their relationship between where they are and where they are going leaving them in a wide range of emotional states</td></tr>
<tr><td>Context</td><td>• The Web
• E-ink display
• Phone (or other portable audio device)</td></tr>
<tr><td>Description</td><td>1. A website that displays your current location and informs the user about the direction and distance from their destination. Can serve multiple users with directions to multiple locations.
2. An electronic display which informs the current user about the direction and distance from their destination. It dynamically changes based on the chosen destination of the current user currently at the device
3. A mobile device which gives the user information audibly.</td></tr>
</table>

Table Q-3 Blapr: Piecing together a signpost

<table>
<tr><th colspan="5">Teasing apart tour guide</th></tr>
<tr><th rowspan="2">Experience (1)</th><th rowspan="2">Surface elements (2)</th><th colspan="2">Experienced effects (3 & 4)</th><th rowspan="2">Distilled experience (5)</th></tr>
<tr><th>Literal (3.1)</th><th>Abstract (3.2)</th></tr>
<tr><td rowspan="2">Helpful fellow who guides you around and shows you where things are</td><td rowspan="2">Movement
Interactive
Context sensitive feedback</td><td><u>Conveying information</u>
Shared experience
<u>Social experience</u></td><td>Comfortable
<u>Informed</u>
<u>Personal Touch</u></td><td rowspan="2">A method of conveying <u>area specific information</u> in a <u>social manner</u> with a <u>personal touch</u>.</td></tr>
<tr><td colspan="2">Remember to note key effects!</td></tr>
</table>

Table Q-4 Blapr: Teasing apart a tour guide

Piecing together tour guide	
Distilled experience	A method of conveying area specific information in a social manner with a personal touch.
Context	• The Web • Audio • Augmented Reality
Description	1. A website that provides you with area specific information that has been provided/conveyed by other users. (Interactive) 2. A non-interactive, pre-recorded audio guide providing area specific information. 3. An application that provides area/user specific information based on input from AR technology.

Table Q-5 Blapr: Piecing together a tour guide

Teasing apart Blapr				
Experience (1)	**Surface elements (2)**	**Experienced effects (3 & 4)**		**Distilled experience (5)**
		Literal (3.1)	**Abstract (3.2)**	
A tool to provide the location of people and places and route from one location to another	Map Sametime Notes plugin Simple input (messages)	Visual information Orientation on the site Ask a question, receive an answer (Sametime) Automation (notes plugin) Available in many places (QR codes)	Personally tailored experience Creates feeling of assurance	A mechanism to provide location data about any person, place or thing and give perspective on the distance and direction of the discovered location.
		Remember to note key effects!		

Table Q-6 Blapr: Teasing apart Blapr

Comparison of results: 'blapr', 'tour guide' and 'signpost'

The TAPT analyses of a tour guide and a signpost have both revealed interesting results whose value and relevance to the design of a product like blapr is not immediately obvious. In performing the same process of deconstruction to the end product, the abstract and literal effects of using either existing solutions or blapr seem to converge.

As an example, a signpost provides a sense of distance and orientation through its shape and the addition of the number of miles to the destination. The shape of a sign tells the person reading it whether it marks the boundary of a town or city (rectangular, no distance measurements) or if it is pointing towards the town (arrow shaped.)

In the same way, blapr shows the location of a destination on a map, providing a sense of distance from the user's current location and a feeling of orientation by showing both points on the same map.

These similarities are useful as a way to gauge the ease of use and quality with which the original brief has been satisfied. The use of TAPT at the start of a project is best felt when deciding the most appropriate format to deliver the solution based on user's previous experience and expectations for the presentation of information.

The 'check' to make sure the end product has delivered on its goals of providing its end users with a product to satisfy their needs in a familiar way is achieved by going through the TAPT process a second time, with the final deliverable as the focus. The data from both processes should correlate, in at least an abstract way, to ensure that the user of the provided system has an experience to match their expectations.

The reason for this is that both the design and evaluation stages are very complementary ,in that each provides a context of understanding for the other. With only the design stage, there is no concrete way to measure the effectiveness of designing a product around the gathered design data. Using just an evaluation stage, the only knowledge gained is the kind of experience that the product creates; there is nothing to say that the experience is the correct one.

Spoken Web

Teasing apart results

Teasing apart Status updates				
Experience (1)	**Surface elements (2)**	**Experienced effects (3 & 4)**		**Distilled experience (5)**
		Literal (3.1)	**Abstract (3.2)**	
You can post a brief description of what you're thinking/doing and see similar descriptions from your friends.	-brief -privacy options -text/photos/links -comments -list -tags -web based -quick	-information feed -networking -conversations -feedback	-curiosity -feel involved -help feel in touch with people -share emotions -secure	Facebook provides a secure environment to share emotions with your friends and have conversations based on what you see in the information feed.
		Remember to note key effects!		

Figure Q-1 Spoken Web: Teasing Apart status updates

Piecing together Status updates	
Distilled experience	
Context	Spoken web
Optional sketches / diagrams / flow charts	
Description	A way to record a short verbal update of your current activity or thoughts, published in the ~~public domain~~ social network of the person. Contacts of the person may record & listen to comments on the update.

(Security to be addressed)

Figure Q-2 Spoken Web: Piecing Together updates

Teasing apart results News feed and

Teasing apart Notifications				
Experience (1)	**Surface elements (2)**	**Experienced effects (3 & 4)**		**Distilled experience (5)**
		Literal (3.1)	**Abstract (3.2)**	
See updates about a group of people's recent activity.	- shared experience - text - photos - links - new friends - conversations - timely - saves time - won't miss things - blacklist	- email - networking - conversation - feedback - starting point - clutter - broadcast - selective	- curiosity - annoyance - informed - inspired - feel closer to people	The news feed broadcasts information to a selective group of people. It's a starting point so you can be informed about your ~~friends~~ friends and feel closer to them.
		Remember to note key effects!		

Figure Q-3 Spoken Web: Teasing Apart notifications

Piecing together News feed & notifications	
Distilled experience	
Context	Spoken web
Optional sketches / diagrams / flow charts	
Description	Accessing statuses, wall posts, & comments on both of those. Updates will only be from friends, not FOAF. The 10 most recent updates are given, & updates are audio links to profiles. Unregistered people, if given the phone number, can access the data set and leave messages. Ability to create blacklists & whitelists.

Figure Q-4 Spoken Web: Piecing Together notifications

Gaming

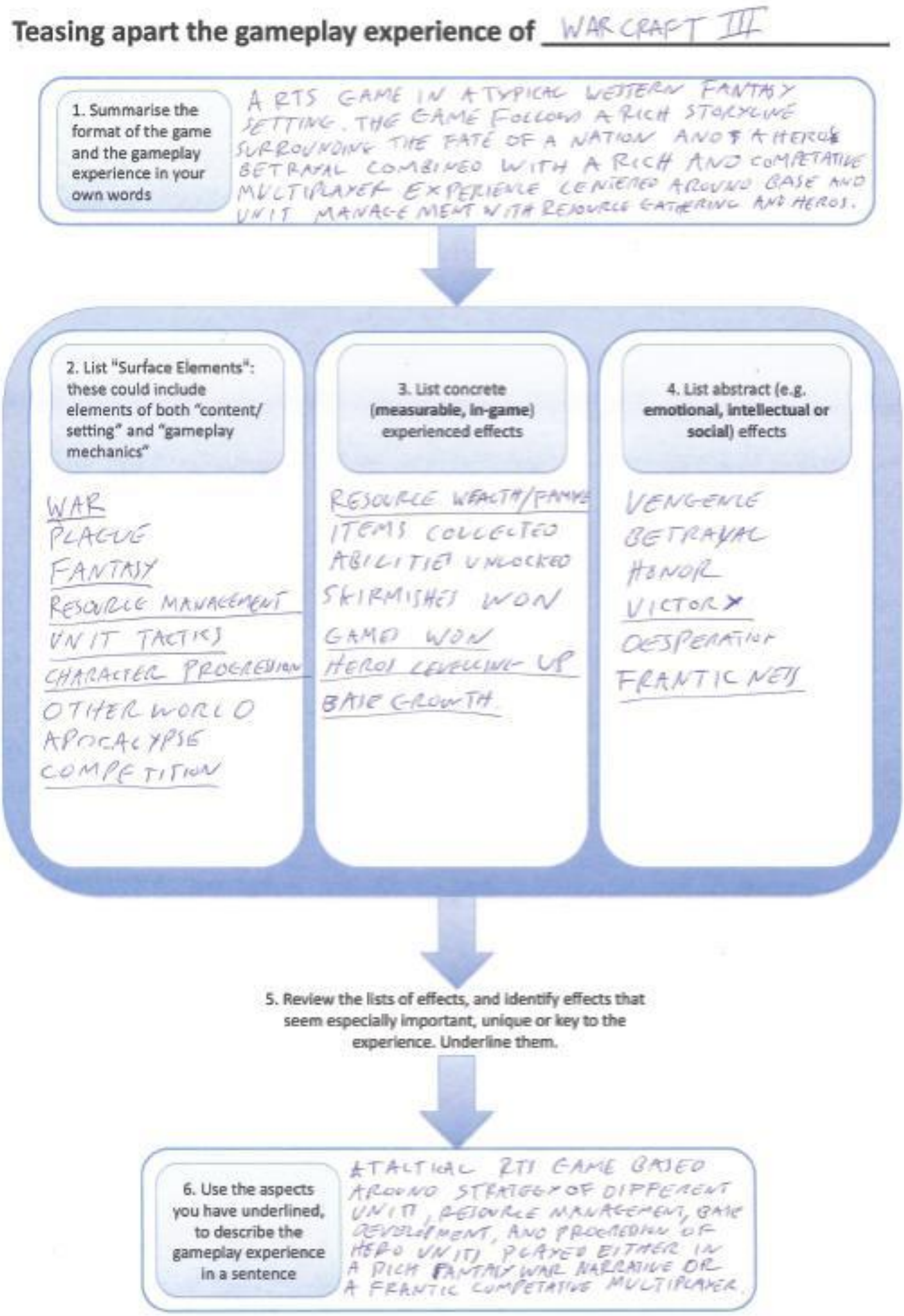


Figure Q-5 Sample TAPT artefact from Gaming case study

Location-based tools

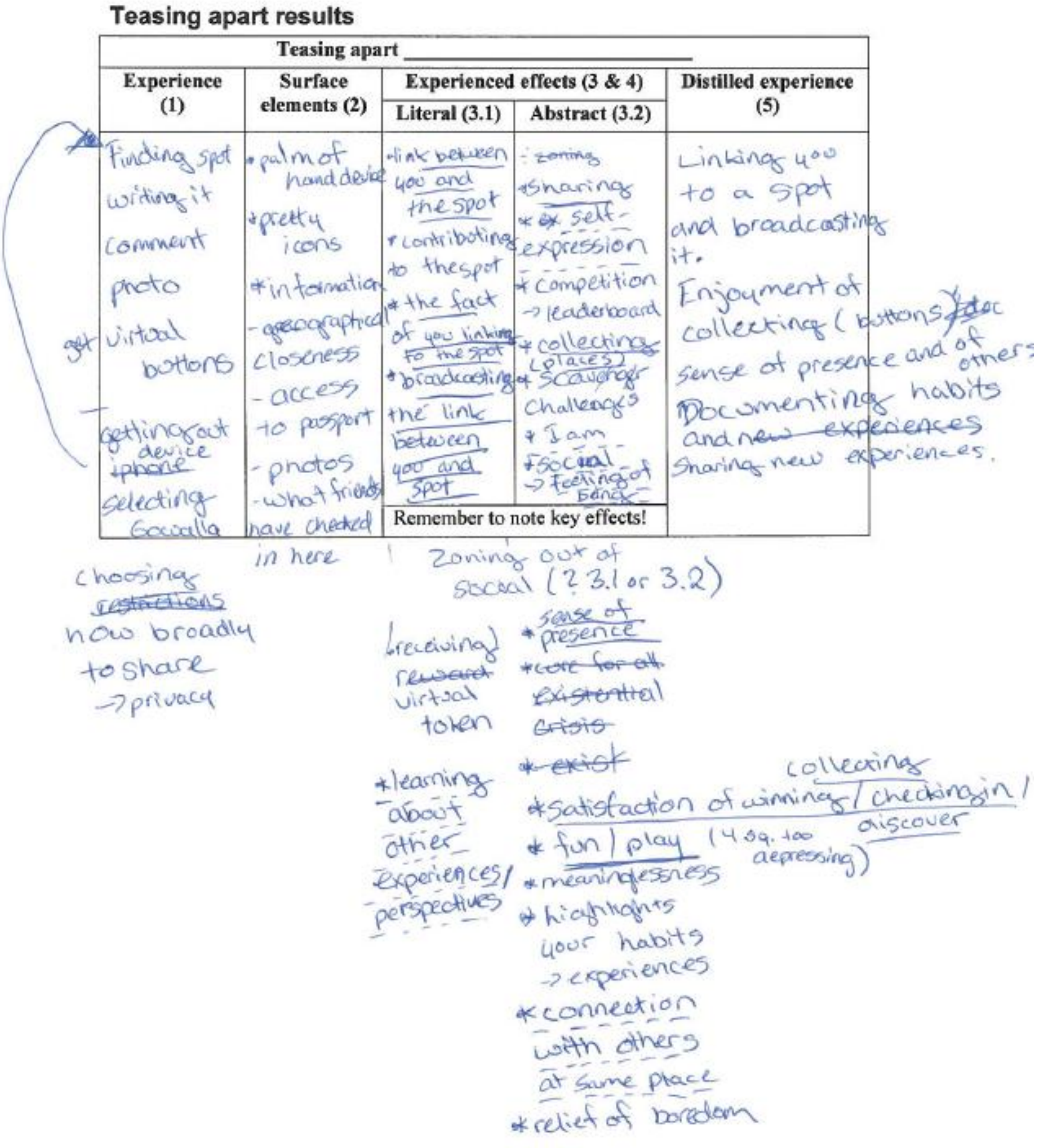


Figure Q-6 Location-based Tools: TAPT analysis of Gowalla

Teasing apart results

Teasing apart ______________				
Experience (1)	**Surface elements (2)**	**Experienced effects (3 & 4)**		**Distilled experience (5)**
		Literal (3.1)	**Abstract (3.2)**	
GPS-based treasure hunt where you use your Phone or GPS device to get to the Point in order to Search for the Cache. When you find the Cache you log it.	• get out and about. • box/container in various Sizes. • urban/rural • GPS • Geocaching.com	• toys/small gifts • Hidden • Search • log	- Share · excitement · inconspicuous (muggles) TFTC! · dissapointment · trade · cooperation	
		Remember to note key effects!		

Try to hide the fact that you are Searching for Something. So the Muggles (People who do not know about geocaching) won't see/understand what the hell it is you are doing. (Secret Society)

TFTC: being able to write this both in the log in the cache and in your post when logging your find. Sort of feeling great about what you have accomplished.

Secret ~~challenges~~ challenge where you go on a treasure hunt Searching for logs

Figure Q-7 Location-based Tools: TAPT analysis of geocaching (first subject)

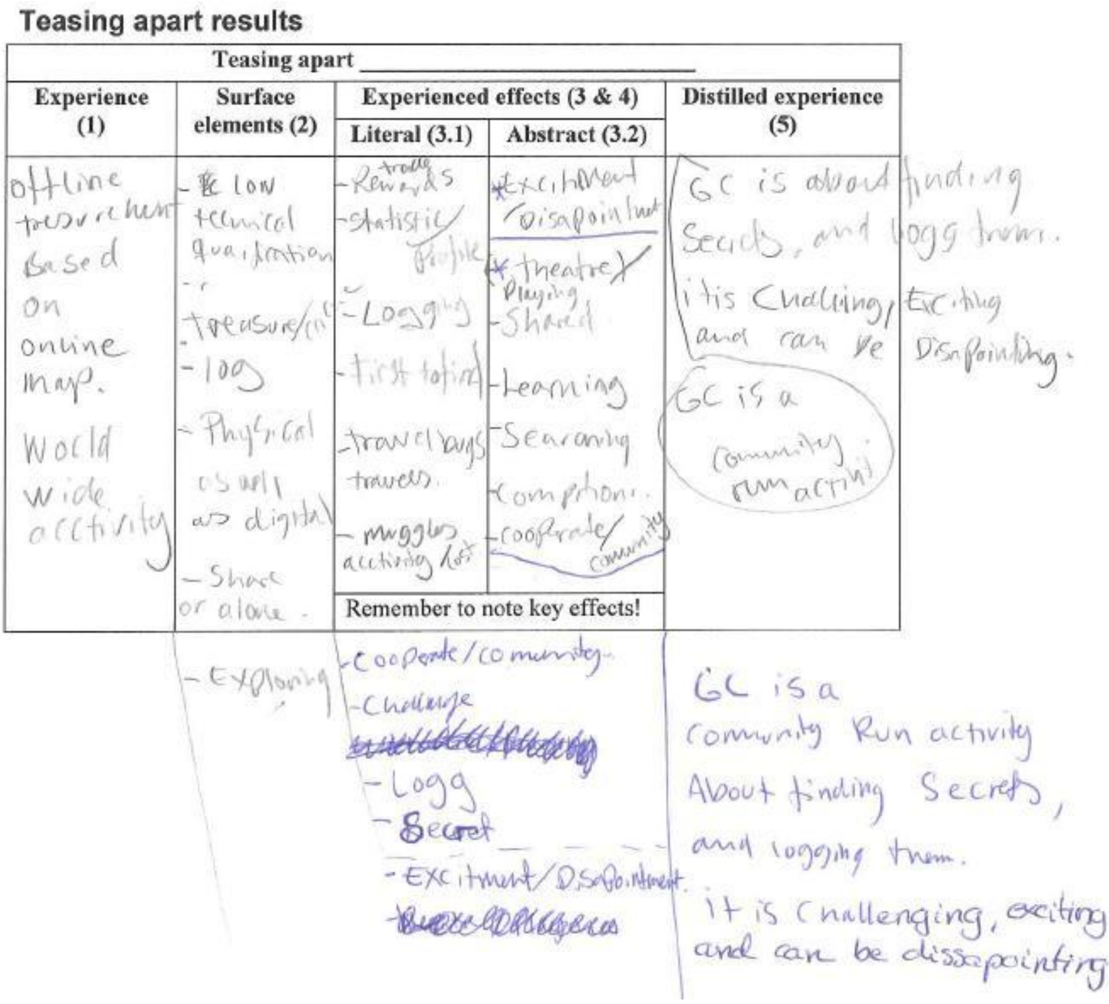


Figure Q-8 Location-based Tools: TAPT analysis of geocaching (second subject)