Human to Computer Trust in Urban Pervasive Computing

Bevan, Chris

Award date:
2011

Awarding institution:
University of Bath

Human to Computer Trust in Urban Pervasive Computing

Christopher Richard Bevan

A thesis submitted for the degree of Doctor of Philosophy
University of Bath
Departments of Computer Science and Psychology
September 2011

Abstract

How people come to trust computing technology is an important factor in the degree to which they come to accept the services that such technologies are able to provide. This is particularly important where the usage of a technology might risk compromising a person's private information, making them vulnerable to technologically mediated attack. Research into interpersonal trust development between people has allowed designers of systems deployed using technologies such as the World Wide Web to successfully modulate a number of human interpersonal trust cues into the computer-mediated communication domain.

Pervasive computing however, describes a significant shift in the ways in which people will come to encounter and use interactive technologies. No longer limited to the confines of the desktop, people can and will be able to perform many of the technological activities they would otherwise do at home in any place and at any time. However, while the services that a pervasive computing infrastructure will be able to provide may be similar to those that people are used to in the traditional world of the 'desktop metaphor', the novel characteristics of pervasive computing mean that many of the trust cues that were previously available to technology users may no longer offer an effective means of helping users to place their trust well.

In this thesis, a mixed methods research plan, involving both laboratory-based and field-based experimental design, was developed to investigate the role of human-computer trust in respect of two novel characteristics of pervasive computing: *service discovery* and *secure ad-hoc device association*. Through leveraging various artefacts in the immediate physical world to support information presented by services in the digital world, this thesis posits that the provision of user-verifiable links between the physical and digital worlds might provide a means of increasing user trust in services whose source they can otherwise not perceive nor verify.

Contents

List of Tables

List of Figures

To Keith and Millie Bevan, and Anne Searchfield.

Great characters all.

0.1 Publications

Sections of work reported in this thesis have previously appeared in peer-reviewed publications.

The study detailed in chapter five was presented as a long paper at the 2009 International conference on Ubiquitous Computing (Ubicomp):

> Kindberg, T., Bevan, C., O'Neill, E., Mitchell, J., Grimmett, J., and Woodgate, D. 2009. Authenticating ubiquitous services: a study of wireless hotspot access. In Proceedings of the 11th international Conference on Ubiquitous Computing (Orlando, Florida, USA, September 30 - October 03, 2009). Ubicomp '09. ACM, New York, NY, 115-124.

A long paper detailing the cyptographic mechanisms of the *Interlock* protocol that formed one of the interactions presented in chapter five was presented at the 2009 IEEE International Conference on Wireless and Mobile Computing (WIMOB):

> Kindberg, T., Mitchell, J., Grimmett, J., Bevan, C., and O'Neill, E. 2009. Authenticating Public Wireless Networks with Physical Evidence. In Proceedings of the 2009 IEEE international Conference on Wireless and Mobile Computing, Networking and Communications (October 12 - 14, 2009). WIMOB. IEEE Computer Society, Washington, DC, 394-399.

The study detailed in chapter seven was presented as a long paper at the 2008 Annual SIGCHI Conference on Human Factors in Computing Systems (CHI):

> Kindberg, T., O'Neill, E., Bevan, C., Kostakos, V., Stanton Fraser, D., and Jay, T. 2008. Measuring trust in wi-fi hotspots. In Proceeding of the Twenty-Sixth Annual SIGCHI Conference on Human Factors in Computing Systems (Florence, Italy, April 05 - 10, 2008). CHI '08. ACM, New York, NY, 173-182.

0.2 Acknowledgements

A great many people have contributed their time, energy, ideas and patience in helping me with the pursuit of this research. First and foremost I would like to thank my supervisors: Eamonn O'Neill, Danaë Stanton Fraser and Tim Kindberg. I could not have wished for a better team to have worked with and learned from. Thank you all for everything you have done for me over the years.

Throughout the course of any protracted research process there are numerous occasions where one finds oneself, to use the Yorkshire vernacular, *on Ilkla Moor baht 'at*. Helping me both to find my hat and keep it on when it mattered most is due in no small part to the encouragement and support of my Cityware colleagues: Tim Jay, James Mitchell, Jim Grimmett, Vassilis Kostakos, Dawn Woodgate, Ava Fatah gen. Schieck and Tim Jones.

Several of the studies described in this thesis were made possible only by the kindness and support of others. To that end, I would like to thank the management and staff at the *Watershed* (Bristol), *Gordon's Cafe* (University College London) & the *Dolce Vita Cafe* (University of Bath) for allowing me to conduct research activities on their premises. I would also like to take this opportunity to thank Jo Reid, Eric Geelhoed and the staff of Hewlett Packard Labs Bristol, for without them I would probably not have found myself writing this thesis at all.

To my family who, despite my many years of study and lack of 'proper jobs', have always kept my back. I can only assure you that, while I still do not entirely know where I will end up, I now at least have a vague notion of where I am heading. Bear with me, I hear the parental gloating rights alone are well worth the investment.

Finally to Freya. May life forever bring you the upmost fuzzy bass-edness. x.

Chapter 1

Introduction

> "The age of ubiquitous[1] computing is here: a computing without computers, where information processing has diffused into everyday life, and virtually disappeared from view. What does this mean to those of us who will be encountering it? How will it transform our lives? And how will we learn to make wise decisions about something so hard to see?"
>
> - Adam Greenfield, *Everyware*[2], 2006 [58].

1.1 Urban pervasive computing

The notion of *pervasive computing*, or 'computing everywhere', describes a significant shift in how and where people will come to encounter and use information technology. In a pervasive computing world, computing power will be freed from the desktop and woven seamlessly into the very fabric of the world around us. Within this vision, services that bear relevance to the current location and context of their users will be made possible through a vast interconnected network of deeply embedded sensors and processors that constantly monitor, process and share information about the world and its inhabitants. At time of writing, the grand theoretical vision of pervasive computing is edging towards an practical everyday reality. Publicly accessible wireless networking infrastructure is rapidly finding ground in many urban areas and, in parallel, the personal computing devices to be found in the pockets and bags of the public *en-masse* are becoming increasingly connected, and increasingly connect-able. Accordingly, *location-based* computing services that fundamentally rely upon mobile and ad-hoc networking communications to provide contextually relevant information

[1] The term *pervasive computing* continues to be used interchangeably with the term *ubiquitous computing* as originally discussed by Weiser (1993) [150]. Though semantic differences between usage of these terms do exist, those differences are not considered as having sufficient enough impact to warrant differentiation within this thesis. This thesis will therefore maintain use of the term *pervasive*, and will consider the terms *ubiquitous* and *pervasive* as being synonymous.

[2] This specific quote was sourced from the website accompanying *Everyware*: http://www.studies-observations.com/everyware/samples.html.

are now becoming a common feature of the urban landscape.

Pervasive systems are sustained by information about the context of their users. Perhaps the most important aspects to generating any reasonable understanding of a users current context, as researchers such as Schilit et al (1994) [130] have discussed are *where [they] are*, *who [they] are with*, and *what resources are nearby*. Thus, as pervasive computing services continue to evolve and mature, there will be an increased and unavoidable requirement for those services to be able to both glean and handle potentially sensitive information about people and the activities in which they engage (Adams & Sasse, 2001 [3]). These characteristics of pervasive computing will increase the need for pervasive system designers to consider the potential impact of their designs upon the privacy and security of their users' data. In parallel, the machinations and requirements of pervasive technologies, particularly with regard to issues of personal data privacy and security will levy a substantive and new requirement upon human-computer trust.

Central to the use of human trust is the identification of *who* or *what* is under consideration (Seigneur & Jenson, 2005 [89]). As the infrastructure that supports pervasive services integrates to such an extent that its workings are no longer visible, it will become increasingly difficult for users to know what systems are functioning and who is offering those systems. As a result, their knowledge and understanding about how such systems might affect their privacy may be severely constrained (Beckwith, 2003 [12]). For all computing systems that require access to personal and / or private user information (be they pervasive or otherwise), the ultimate user acceptance of those systems will be largely dependent upon the level to which those systems are considered *trustworthy*. As will be discussed, though *trustworthiness* is not trust, it is a significant antecedent factor of trust. For new service vendors, this distinction offers opportunities that are of critical importance. Designers must now, more than ever, convey to users that their systems are trustworthy before their users are able to prove it through positive experience (Koufaris & Hampton-Sosa, 2004[79]). The degree to which pervasive systems can successfully convey their trustworthiness to potential users is the central theme of this thesis, and the nature of the problem is twofold (Kindberg et al, 2008 [74]):

- Users may mistakenly invest their trust in a malevolent system that was mistakenly considered as being trustworthy (thus inviting attack).
- Users may fail to invest their trust in a bona-fide service (and thus miss out on its benefits) because the service failed to communicate its trustworthiness effectively.

1.2 Human-computer trust in urban pervasive computing

Interactive and interconnected technologies such as the World Wide Web (WWW) have revolutionised the way in which people are able to communicate with one another and conduct their daily business activities. In tandem with the benefits afforded by these technologies

however, there exist many opportunities for criminals to subvert their use for malevolent ends.

That users come to trust the services provided by pervasive technologies is a critical factor in the degree to which such services will ultimately succeed and flourish. Designers of WWW based services such as electronic 'e'commerce and online consumer banking services have long recognised the importance of communicating the trustworthiness of their systems to their customers (Lee & Turban, 2001 [82], Yousafzai et al, 2003 [153]). However, to simply convey trustworthiness alone is not enough. The positive effects of trust can only be found where user trust is 'well-placed' (Riegelsberger, 2005 [119]). Systems designers must consider the effect of a false-negative trust investment (i.e. mistakenly accepting a malevolent system) as being a substantially less desirable outcome than a false-positive investment (i.e. mistakenly rejecting a bona-fide service). Indeed, to merely uncover and exploit factors that increase perceptions of trustworthiness at a superficial level can instead result in the undermining of user trust. As Riegelsberger (2005 [119]) noted, the designer's focus should be directed not just upon increasing user perceptions of trustworthiness per se, but also (and substantially so) on increasing their ability to discriminate between what is worthy of their trust and what is not.

From the perspective of computer science, solving the problems of human-computer trust has been considered for a long time to be synonymous with solving the problem of security. Within such a worldview, the panacea would ultimately be found in the development of ever-stronger cryptographic measures (Viega et al, 2003 [149]). With particular respect to the advent and subsequent development of the WWW, equating trust and cryptographic security has served information and communication technology (ICT) service developers well. To date, significant improvements have been, and continue to be made in data encryption technologies that have served to allay very real threats to private / sensitive data. Stronger encryption and increasingly novel methods of establishing user authentication have now made high-risk activities such as administering ones own bank account via a still largely unregulated Internet a well established and accepted practice. However, how (and indeed whether) such techniques are sufficient when transferred to similarly risky activities within the pervasive domain is unclear. There is a marked difference between how pervasive computing services communicate with users and how this was achieved with traditional forms of networking. In times past for example, access to computing-infrastructure based services was generally achieved through a physical connection between the user device and the host network, i.e. through an ethernet cable and a network 'port'. From the standpoint of trust, the physical connection to a network through a 'port-in-the-wall' of a trusted establishment might allow a user to be reasonably confident that they were then connected to the network / service that they expected. However, with wireless networking now becoming the norm, the visibly verifiable end-to-end link that the cable was once able to provide is now becoming less and less available.

Further, the distributed, wireless and invisible nature of pervasive service provision creates a number of new routes for malevolent activity to which purely technical features such as cryptographic security may no longer be an effective remedy (Butler et al, 2003 [19]). It is, for example, perfectly feasible within a pervasive computing scenario for a user to interact over a 100 percent secure connection with a bogus service provider (Grazioli & Jarvenpaa, 2000 [57]). The lack of a visible source of origin for wirelessly deployed services

may thus place demands upon human-computer trust that are above what would be expected from an analogous service delivered through more traditional modes of communication. Furthermore, as the infrastructure that drives the pervasive computing world continues to fade into the background, so too will many of the cues that may have helped users to assess the authenticity, intention and motives of previously unencountered computing services in the past.

1.3 Background and scope of the thesis

The research presented in this thesis was conducted within the UK Engineering and Physical Sciences Research Council funded research project *Cityware: urban design & pervasive systems*, grant EP/C547683/1. Specifically, this research formed part of a sub-component (work package) of *Cityware* that sought to investigate the role of *security, privacy and trust* as they related to urban pervasive computing.

1.3.1 Relationship with the Cityware research project

Cityware (2006-2009) was a three year EPSRC funded research project created to "develop theory, principles, tools and techniques for the design, implementation and evaluation of city-scale pervasive systems as integral facets of the urban landscape."[3] A multidisciplinary research project, *Cityware* involved aspects of architecture and urban design, human-computer interaction and computer science. Project partners involved in the project included The University of Bath, Imperial College London, University College London, IBM United Kingdom Ltd, Vodafone, Bath and North East Somerset Council, Hewlett Packard Labs U.K. and Nokia.

Within *Cityware*, the author was engaged as a member of a team assigned to a specific work package: *security, privacy and trust*. The general aim of this work package was to develop an understanding of the mechanisms needed to defend users against threats to their privacy and security when they interact with services in urban environments. Within this work package, the author would work closely with Tim Kindberg, spending at least one day of his working week at Hewlett Packard Labs in Bristol.

A conceptual / theoretical aim of the *security, privacy and trust* work package, and the aim to which this thesis was developed to address, was to investigate the role and impact of human-computer trust when encountering and using pervasive digital services for the first time. Thus, the primary aim of the research reported here was to develop an understanding of the decision making process involved when people are obliged to invest their trust in order to (voluntarily) use a pervasive computing service of which they have no prior experience. The research presented in the thesis will therefore focus upon investments of a specific type of trust: *initial situational trust*. These are instances where a spontaneous trust-based

[3]This specific quote was sourced from the website accompanying *Cityware*: http://www.cityware.org.uk.

relationship is made manifest ad-hoc between a human and a system that have no prior experience of one another.

Given the author's role as a member of a team of researchers, both with respect to the specific work package previously described, and to the *Cityware* project in general, there was a requirement from the outset to incorporate a degree of flexibility in how the research reported in this thesis was conducted. This feature was made particularly salient in respect to the empirical component of the thesis. For example, while the author's remit (in terms of how he came to understand and operationalise human-computer trust) was fairly wide, it was considered critical to *Cityware* that the work of the thesis would need to be directly relevant enough to be incorporated into deliverable security protocols. These protocols would also be worked on by members of the project that were not necessarily directly working on the same work package.

1.3.2 Thesis scope: Pervasive situated services

The research presented in the thesis will focus upon a particular manifestation of pervasive computing that the author will refer to as *pervasive situated services*. *Situated* computing has been described by Hull et al (1997) [64] as "the ability of computing devices to detect, interpret and respond to aspects of the user's local environment". In the context of this thesis, the term *Pervasive situated services* (e.g. Zambonelli, 2011 [156], Castelli et al, 2007 [21]) will refer to digital services and supporting infrastructure that are embedded within, and directly referent to, particular places and contexts. This is in contrast to other uses of the term *situated* that have been used elsewhere in the literature; an example being Lucy Suchman's *[Plans and] situated actions* (Suchman, 1987 [144]).

A core feature of pervasive situated services as described here is that they must reflect some aspect of the immediate physical environment in which they are encountered. The value of such a capability allows for the creation of ICT services that are able to both perceive and interact with information that is gleaned from the immediate physical and social world of the user (Zambonelli, 2011 [156]). In achieving this, pervasive situated services are able to deliver information to users that is naturally highly contextually relevant. An example of a pervasive situated service might be digital content that refers to nearby physical artefacts, such as might be found in an environment such as a museum. In such a scenario, a pervasive situated service might be deployed within the museum building that would offer (and deliver) additional information about particular exhibits to a visitor's mobile device. Such a facility might well be provided as a value added service by the owners of the museum and would only be available to museum patrons who are physically present in the museum building.

When discussing the notion of 'situated-ness', the degree to which infrastructure and services are attached (and limited) to the immediate environment are key measures. In terms of infrastructure, a large display, that fixed / bolted to a wall, would therefore be considered (in the context of this thesis) to be highly situated. Similarly, digital service provision delivered through proximate / medium-range wireless digital radio signals (e.g. using technologies such as WiFi, 'Bluetooth' and 'Near Field Communication') can also be considered as being 'situated', as receipt of such signals could only be achieved through being in close physical

Figure 1.1: *Free Wifi here*: Public access wireless Internet services are becoming a common feature of the urban landscape. This example was observed being displayed on the door of a cafe in Bristol.

proximity to the transmission source. However, it must also be considered that a given situated service itself need not exist in only one place at one time; they are free to manifest themselves on a more ad-hoc basis, perhaps forming spontaneously from an ecology of mobile devices that simply come to be co-located in a shared space and time.

Within this thesis, the pervasive situated service that will be examined most directly are wireless Internet gateway services, commonly referred to as 'WiFi hotspots' (figure 1.1). An increasingly common feature of the urban landscape, WiFi hotspots are frequently encountered in public spaces such as cafes, libraries and public transport hubs.

Alternative technologies such as *Bluetooth* and *Near Field Communication* (NFC) were also considered as candidates for examination during the early stages of the research programme. However, the decision to use WiFi hotspots as a focus for this research (and in particular the empirical component of the thesis) was made over these other technologies for several reasons. As a wireless ICT service of limited range that can be made available in particular places and times, it was felt that WiFi hotspots possessed most (if not all) of the core characteristics of a pervasive situated service as previously described. Further, as the empirical component of the thesis would likely involve the deployment of a working pervasive situated service in the field, the author felt that certain characteristics of WiFi hotspots made them an excellent candidate as a research device. Connecting to the Internet wirelessly was a process that was already quite familiar to the public - if not in public space, then certainly within the home. This was considered (in 2006-2007) to be somewhat less the case with ad-hoc connections using Bluetooth, and lesser still with Near Field Communication. Lastly, the use of WiFi presented the opportunity to utilise interactive content delivered through a web browser as an experimental device. As the author already had considerable professional experience designing and developing interactive web content, this characteristic was considered ideal.

1.4 Justification for the research and research contributions

The novel nature of pervasive systems invite investigation of two major issues with pertinence to human-computer trust. These issues are:

1. The wireless / invisible nature of pervasive systems constrain user ability to ascertain the source and intention of the services they provide.
2. User trust is often obligated by the requirement to surrender personal information prior to engagement with such systems.

A requirement for human-computer trust with pervasive computing services can be found at two levels: computer-computer trust (in order to maximise data transport security and efficiency) and human-computer trust (in order to maximise user acceptance). While significant research efforts have been, and continue to be made on the former (e.g. Kagal et al, 2001 [70], Satyanarayanan, 2001 [129]), research into human-computer trust that specifically tackles services deployed in the pervasive domain is currently lacking. Little is currently known about how people understand pervasive services, and less still in how they might consider and assess them in terms of their trustworthiness. The research literature that does investigate how mobile device-based trust-relevant interactions (such as e-commerce-based monetary transactions) differ from established models remains scarce (Kindberg et al, 2004 [75]).

Such that pervasive systems require the ability to ascertain (to some degree) whom they are to deal with, the requirement of trust on the part of the user is often mandatory. If a person wishes to obtain the benefits that such systems purport to provide, they must invest their trust in those systems; they have little in the way of alternative options available to them, other than to waive the benefits of their use. Despite considerable progress however, consumer ICT services that are based upon a pervasive infrastructure are still very much in their infancy. Whether existing research and techniques to maximise trustworthiness and well-placed user trust can generalise from the static computing systems for which they were designed to computing on the pervasive level remains unclear. The novel contributions of the thesis are thus twofold:

1. **Theoretical**: To develop a better understanding of human-computer trust as it relates to nascent pervasive technologies.
2. **Practical**: To develop and evaluate means by which the trustworthiness of benevolent urban pervasive services can be conveyed to users in a way that facilitates their discrimination from similar (and possibly malevolent) services.

1.5 Research questions and hypotheses

The main aim of the thesis is to develop an understanding of how people come to trust emergent pervasive technologies, and how they might use evidence gleaned from their context and environment in order to facilitate this. In doing so, the research presented in the thesis sought to address how people assess the trustworthiness of such systems, and how designers of such systems might incorporate those processes within their designs.

The primary prediction of this thesis is stated thus: *with no a-priori knowledge of the identity or origin of a given situated service, the decision on whether or not to trust (and thus elect to use) that service will be informed in substantial part by the context and environment in which the situation occurs.*

To be clear, the goal of the present thesis was not to provide solutions for improving the security of systems per se. Instead, the present project sought to identify and investigate the efficacy of specific evidential cues that a user can both recognise and understand in order to help ensure that their trust, if invested, was sufficiently 'well-placed' (Riegelsberger, 2005 [119]). The overarching research questions of the present thesis are therefore presented thus:

1. **RQ1:** How is human-human interpersonal trust understood to operate, and how has this been modulated for use in human-computing applications where some degree of trust (on the part of the human) is obligated?
2. **RQ2:** To what extent do people currently understand the threats and risks associated with the use of ICT services such as the WWW? To what extent does this understanding translate to pervasive / situated computing services?
3. **RQ3:** Which aspects of the situational context are considered important to users when they attempt to evaluate the trustworthiness of a pervasive situated service?
4. **RQ4:** How do people utilize the aspects identified in **RQ3** to make decisions about whether or not to invest their trust in a given situated service?

1.6 Overview of the thesis

In order to investigate the role and mechanics of trust in a pervasive situated service scenario, a mixed-methods research plan was devised by the researcher and deployed across four separate studies. Each study is described over a single chapter. Across the four studies, data was collected using a combination of laboratory and field-based work, utilising a range of quantitative and qualitative measures.

Chapter two initiates a discussion of trust in pervasive situated services through a critical discussion of research relating to human-human, human-computer and finally human-pervasive computer trust (addressing **RQ1**). In **Chapter three**, through the development and use of a questionnaire, an initial investigation was conducted by the researcher to probe

the extent to which current users of technology consider their privacy and security in their current use of ICT, and the degree to which these considerations might transfer to less familiar pervasive computing scenarios (addressing **RQ2**).

The results of the questionnaire were then used to inform the basis of the development of an empirical study programme that would seek to explore issues of trust as they relate to pervasive situated services. **Chapter four** describes the design of an empirical study programme and identifies the goals of that programme.

The empirical component of the thesis then begins in **chapter five**. To develop a deeper understanding of user perceptions of privacy, security and trust with regards to specific instances of pervasive situated service use, chapters five and six describe two studies that utilised a semi-lab-based experimental design to allow the researcher to gather qualitative data from participants directly. Both studies measured participants' stated *intention-to-trust* based upon their evaluation of contextually based evidence relating to the potential 'genuineness' of a public WiFi 'hotspot' (addressing **RQ's 3 & 4**). **Chapter five** describes an experiment that was designed to evaluate ways in which evidence made available in the immediate physical surroundings of the participants could serve as a indicator of the trustworthiness of a public access WiFi 'hotspot'. Within the experiment, several co-located situated services were made available, and each utilised various physical and virtual artefacts in the surrounding environment of the participants in order to present themselves as being genuine. **Chapter six**, expanding upon the findings of chapter five describes a semi-lab based experiment in which the presence and behaviours of *other people present in the immediate vicinity of the participants* was utilised as a means of providing evidence about the genuineness of a 'WiFi hotspot'. Concluding the experimental studies **chapter seven** presents a novel 'unattended' field experiment that was designed to engender true risk and to measure actual trust investment behaviour in response to a real-world situated service usage scenario.

Finally, **chapter eight** summarises the contributions and findings of the thesis and provides recommendations to researchers and design practitioners.

Chapter 2

Literature Review

> "Trust is an investment by the trustor which will pay dividends only in the event that the trustee behaves appropriately."
>
> - Philip Pettit, *The Cunning of Trust*, 1995: 216 [113].

2.1 Chapter overview

This thesis is focussed upon how 'well-placed' human-computer trust, specifically with regard to a relatively novel variety of pervasive computing systems called *situated* services, can be both developed and supported by designers of such services. This chapter provides the conceptual background to the thesis.

Such is the range of academic disciplines to which trust is considered an important topic for research, a comprehensive cross disciplinary examination of the wealth of literature relating to trust would be a formidable task. As such, the author chose not to discuss the contribution of any particular discipline in turn. Instead, this review will discuss aspects of trust as they have coalesced and emerged from a range of selected disciplines as a whole. The principle source of the research reported in this chapter is the field of human-computer interaction (HCI), a discipline that draws heavily on research generated in psychology, sociology and computer science. However, literature from the fields of management / political science and economics will also be reported where considered as being particularly pertinent to the ultimate topic of human-computer trust.

The chapter begins by introducing the function and role of trust as a device for facilitating asynchronous exchange between humans. The term *trust* is then delineated from the term *trustworthiness*, and theories as to its mechanism are presented and discussed. The chapter will then move on to discuss the prerequisite factors that are necessary in all trust-based interactions, and in doing so separate trust conceptually from other risk-taking behaviours such as lottery-based gambling. The development of trusting attitudes within individuals

is then presented, starting with the development of general trusting attitudes (*propensity to trust*) and then to more context-specific instances of trust (*situational trust*). Finally, research efforts directed at the phenomenon of *initial-situational trust*, where a high-trust relationship is formed spontaneously without prior experience are presented and discussed.

In the final sections of the chapter, various methods by which humans seek to establish the trustworthiness of others as a response to specific situations are presented. This discussion begins with an examination of how the process of assessing trustworthiness has been understood to occur between people face-to-face, before moving on to how designers of ICT systems have sought to modulate these methods into the digital-interactive domain (addressing **RQ1**). The chapter then concludes with a discussion of research investigations into human-computer trust in respect of nascent pervasive computing scenarios, with a particular focus on the new trust-related problems they present, such as service discovery and user verification (addressing **RQ2**).

2.2 Introduction: The function and relevance of trust

As social beings, we are often faced with the decision of whether or not to commit to actions of which success is dependent upon the free choices of others. Trust, "*a psychological state comprising the intention to accept vulnerability based upon positive expectations of the intention or behaviour of another*" (Rousseau et al, 1998 [125]) is a common strategy that is used by individuals and groups to mitigate uncertainties about future events.

As so much of the social human condition involves reciprocal value exchange, trust is generally considered to be an essential part of social and individual human life. Unbound by culture or nationality, its use has been offered as an example of an economic primitive (e.g. Fismann & Khanna, 1999 [44], Zak and Knack, 1998 [155]) and as a socio-cultural norm (Gouldner, 1960 [55]). As the performance of actions based upon trust negates the need for explicit external control structures (e.g. binding legal arrangements), using trust can save transaction costs in terms of time, energy expenditure and resources (Zak and Knack, 1998 [155], Delhey & Newton, 2003 [32], Sitkin and Roth, 1993 [138], Uslaner, 2002 [147]). Such is its usefulness, gaining insight into the means by which high-trust relationships can be successfully formed and fostered has afforded considerable attention from a variety of academic disciplines, including psychology, sociology, and the management / political sciences (Mollering, 2001 [106], Lee & Turban, 2001 [82]).

2.2.1 Initial positioning and definitions for terms to be used

There are a number of terms that relate to trust that will used throughout the thesis. Initial working definitions for those terms are presented thus:

Trustor: The principle actor, to whom the decision to invest trust (or not) ultimately resides.

Trustee: The actor to whom trust is to be potentially invested.

Trustworthiness: A subjective assessment of the qualities of a trustee by the trustor. An internally generated measure of the potential likelihood that a trust interaction will complete as the trustor expects.

To trust: The intentional and voluntary decision made by a trustor to invest their trust in a specific trustee. The act of investing something of value *to* a trustee, based upon the expectation of a mutually agreed future benefit that is to be received *from* the trustee.

To distrust: The intentional and explicit decision for a trustor not to invest trust in a potential trustee, where there had existed an opportunity to have invested trust.

To mistrust: The past-tense result of the act of a trustee reneging upon a mutually agreed trust-based relationship with a trustor. The consequence of an unsuccessful investiture of trust.

2.3 What is trust?

The term *trust* can be used to describe several phenomena, including a type of personality trait and a type of social structure (McKnight and Chervany, 2001 [100]). Though permissible both as a noun and a verb, in its broadest sense trust is, as the sociologist Georg Simmel [137] describes, "*an individual hypothesis that is strong enough to serve as a basis for practical conduct*". Trust can be used to facilitate and lubricate a diverse range of social interactions (Stolle, 1998 [143]) yet trust itself is internal, and its presence can be observed externally only by its effects "*trust is not an objective property of an entity; rather it is a subjective degree of beliefs about an entity.*" Abdul-Rahman & Hailes (2000) [1].

2.3.1 Separating the terms *trust* and *trustworthiness*

Though strongly related, the concepts of '*trust*' and '*trustworthiness*' are both functionally and conceptually discrete. *Trust*, as Shultz (2006) [131] observes, emanates from the trustor, whereas *trustworthiness* is referent to characteristics of the trustee. Adequate delineation of these terms is a critical and oft-overlooked issue (e.g. Hardin, 1993 [60]).

Trustworthiness is a purely subjective value that is afforded to trustee by trustor based upon the level and type of trust that is required within a specific context (Corritore et al, 2003 [27]). The trustworthiness value of a trustee can be decomposed into two basic categories of belief (Doney & Cannon, 1997 [39], Tan and Thoen, 2001 [145], Castelfranchi and Falcone, 2001 [43], Nooteboom, 2005 [110], Deutsch, 1958 [34]):

Capability / Competence Belief: The belief that the trustee *is able to* perform the trusted action to the best of their known competence.

Intention Belief: The belief that the trustee *is motivated and intends to* perform the trusted action to the expectations of the trustor. Various psychological and sociological traits have been posited as being critical in the assessment of intentional beliefs related to trust, the most common being perceptions of *benevolence* and *integrity* (Mcknight & Chervany, 2001 [100]).

2.3.2 The mechanism of trust

Despite widespread, intuitive and often unconscious use, an all-encompassing definition of the precise nature of trust itself continues to prove elusive (Schultz, 2006 [131]). Despite over a century of research, there remains no commonly accepted theory of its mechanism. The ultimate nature of trust, as much as it is discussed at all, is most commonly discussed in terms of being an irreducible type of faith that has been described as "*nothing to do with knowledge, [but is] both less and more than knowledge*" (Mollering, 2001 [106]). More recently, trust has been represented as the combination of faith and weak inductive knowledge (Mollering, 2001 [106]) that is capable of forming a basis for confidence about an ultimately unknowable future event.

As Morton Deutsch (1958) [34] observed, "*if people were omniscient, all actions could be taken with absolute certainty [and] there would be no need for trust to exist*". Thus, despite having no firm notion of its ultimate mechanism, most authors tend to agree that the principle solution that trust presents is to close a rational cognitive gulf that exists between uncertainty and an absolute (and thus impossible) knowledge of the future (Simmel & Wolff, 1964 [137]). As the cognitive demands involved in the hypothetical modelling of the full range of the potential actions of another are often unworkably vast (Egger, 2001 [41], Good, 2000 [54]), using trust to accept a degree of uncertainty as to the future serves a useful means of allowing pure rationality to be temporarily short-circuited. This process, where cognition is allowed to temporarily satisfy itself into considering an unknowable outcome as being certain, has been described by Luhmann (1979) [85] as a *cognitive leap*. The *cognitive leap* refers to a level of reasoning that "*overdraws on the cognitive base*" of pure reason, effectively filling a gap between the limits of reason and the uncertainty of pure faith. A functionally similar concept termed *suspension* has recently been offered by Guido Mollering, where *suspension* is described as acting as a mediator between a base of 'good reasons' to invest in an uncertain decision in order to receive the benefit, and a momentary certainty that allows trust to enact: "*Trust is the mechanism that 'brackets out' uncertainty and ignorance, thus making interpretive knowledge momentarily 'certain' and enabling the leap to favourable (or unfavourable) expectation*" (Mollering, 2001 [106]).

2.3.3 Prerequisites of trust: *uncertainty* and *risk*

> "Trust is the expectation that one will find what is expected rather than what is feared".
>
> - Morton Deutsch, *The resolution of conflict: constructive and destructive processes*, 1973 [35].

The use of trust in interpersonal relationships is linked inexorably to a group of well-established social norms, most directly *reciprocity*, the tendency to respond to one another in kind, and *commitment*, the keeping of one's promises (Cialdini, 2000 [26]). To trust is to have a confident, positive expectation about the outcome of an interaction that is conducted within a situation of risk (Schultz, 2006 [131]).

Within all arrangements that are based on trust, the agreement that is made between the parties involved is ultimately unbound. Indeed, for trust to exist at all, it is crucial that all parties involved are aware of this fact. Two conditions are thus always present in any situation that can feasibly support the investment of trust: *uncertainty* about the future (e.g. Sabel, 1993 [127], Lewicki & Bunker, 1997 [83], Deutsch, 1958 [34]) and *risk* (e.g. Mayer et al (1995) [91]) that the trustee will choose to behave opportunistically. The terms *risk* and *uncertainty* are sometimes incorrectly considered as being synonymous (Riegelsberger et al, 2005 [120]). To be clear, uncertainty refers to situations where adverse effects are possible, but no probabilities are known (or can be known), whereas risk is the possibility of adverse effects with a known (and thus ultimately avoidable) probability. Such is their importance of their contribution to reasoning about trust, it is worth considering both terms in more detail.

Uncertainty is a state of limited knowledge about a future that is unknowable, but where it is perceived that there are several potential outcomes to a given action. As the source of most potential outcomes are at the whim of the trustee, the core source of uncertainty as it relates to trust is directly attributed to the inability of the trustor to attain complete control over the interaction. To this end, it is useful to consider Luhmann's (1979) [85] assertion that trust is a device for "*coping with the freedom of others*".

Risk refers to the degree of probability that at least one potential outcome of a proposed action would yield undesired effects. Adams (1995) [4] suggests that risk can be measured in terms of the compound of probability and the magnitude of the adverse outcomes possible in the event of a negative conclusion. As it relates to trust, risk is directly referent to the knowledge that it is within the capabilities of the trustee to fail to honour their side of the agreement.

Asides the notion that to renege upon a trust investment is generally considered to be a morally dubious act (Good, 2000 [54]), trust offers little in the way of an explicit deterrent from opportunism[1]. As Hardin (1993) [60] notes, "*trust involves giving discretion to another to affect one's interests. This move is inherently subject to the risk that the other will abuse the power of discretion*". Indeed, there has existed a general axiom, particularly within traditional models of economics, that much human behaviour is rooted in individualistic gain maximisation (as researchers such as James (2002) [66] point out, as far as human behaviour goes "*there is a stable preference for more rather than less*"). How a trustor perceives a trustee as possessing freedom of opportunity is thus key. As the freedom by which a person is able to choose their actions increases, the incentive that is present to maximise their self-gain raises also. "*Trust is related to the fact that agents have a degree of freedom to disappoint our expectations, and indeed, trust becomes increasingly more salient for our decisions and actions the larger the feasible set of alternatives open to others*" (Gambetta, 1988 [50]). Further, as Berg et al (1995) [14] observed, "*a fundamental assumption in*

[1] *Opportunism* is defined by Williamson (1985) [151] as *"self interest seeking with guile"*.

economics is that individuals act in their own self interest [...] Behaviour that deviates from self interest is viewed as irrational".

It is generally agreed that notions of trust and risk are dissociable (Deriaz, 2006 [33], Mayer et al, 1995 [91], Deutsch, 1958 [34]), yet it is important to note that trust is not involved in all risk taking behaviour. At a purely functional level, to base a decision on trust appears similar to decisions based on gambling. However, trust and gambling differ importantly in terms of their antecedents and potential consequences. The fundamental difference between trust and gambling is found within notions of perceived control. To invest trust, it is critical that the level of uncertainty (with regards to the eventual outcome) is perceived by the trustor to be within their ability to affect, change or control in some way. Any true investment of trust thus demands that the trustor believes that the trustee has the ability, should they so wish, to behave opportunistically. Conversely, a pure gamble is based on a system where the outcome cannot be influenced in any way by the trustor. Using roulette as an example, we cannot invest our trust in the wheel, for the wheel cannot choose of its own volition to honour our expectations (Shelat & Egger, 2002 [133]). A further useful distinction between trust and gambling is offered by Deutsch (1958) [34]. A person is considered to be gambling only when he perceives "*that his potential gains from taking the risk are greater than his potential losses*". Deutsch illustrates this distinction using the scenario of a mother entrusting the care of her child to a babysitter. If the trust that the mother places in her babysitter to care for her child is fulfilled, her payoff (i.e. the ability to enjoy a night out free of the care of the child) is disproportional to the negative outcome of an unfulfilled trust (e.g. her child coming to harm). However, should the outcome prove synonymous with the mother's positive expectations, to invest trust in this situation would be more advantageous than to not invest trust.

2.4 The development of trusting attitudes in the individual

Returning briefly to the theory of the Sociologist Niklas Luhmann (as presented earlier in this chapter), trust can be thought of as existing due to the complexity of the world, and how we as individuals seek to reduce that complexity to a level that our finite cognitive resources can effectively manage. "*The world presents itself (to any thoughtful person) as unmanageably complex. Trust serves to reduce this complexity with cognitive, emotional and moral expectations that some things will remain as they are or ought to be*" (Luhmann, 1979 [85]).

Most authors agree that general attitudes to trust (e.g. the degree to which a person might agree with the idea that "it is generally better to trust than not to trust") are formed early in life, and then refined over time through life experience. Within the literature, the development of trusting attitudes have been discussed as occurring over a number of stages. The following sections present and examine these stages in turn, beginning with generalised trust attitudes before narrowing down to individual instances of trust behaviour with specific trustees.

2.4.1 *Basic trust* and the development of an individual's *propensity to trust*

Trust is not exclusively found within interpersonal relationships between humans. Trust also serves as a useful cognitive heuristic for a myriad of other general day to day events. At its most pervasive level, it is suggested that there exists for everyone a degree of *basic trust* (Erikson, 1950 [42], Giddens, 1991 [52]). *Basic trust* is the unconscious and constant acceptance of the unavoidable degrees of risk and uncertainty that invade every waking moment of our lives. As Giddens illustrates, basic trust provides a '*screening off*' device: a "*protective cocoon that permits people to get on with the affairs of everyday life*" (Giddens, 1991 [52]).

Beyond and in parallel to basic trust, all people also have a self-defined level to which they trust other people and institutions generally. This is often described as their *propensity to trust*[2]. Psychologists, focussing on individual personality tend to discuss *propensity to trust* as an individual property associated with core personality traits, individual characteristics and individual social / demographic features. Advocates of this view include Erikson (1950) [42], Allport (1961) [7], Cattell (1965) [22] and more recently Uslaner (2000) [146]. Though arguments continue that some level of propensity to trust may be an innate mechanism borne of evolution and present at birth, most authors agree that most of a person's general propensity to trust is a learned behaviour from early childhood that is subsequently revised and refined throughout adulthood. McKnight and Chervany (1998) [101] discuss propensity to trust (or trusting disposition) as being comprised of two levels of subjective belief about people in general. These beliefs are described as a person's *faith in humanity* and their *trusting stance*.

Trusting stance: The belief that, regardless of past experiences of reliability, it is better to deal with people as if they are well-meaning and reliable.

Faith in humanity: A person's level of belief that others are typically well meaning.

Obtaining a measure of general propensity to trust within a population is of interest to social scientists as it offers a useful indicator to current levels of societal cohesion and social capital (e.g. Putnam, 1993 [116]). General levels of propensity to trust have been measured using attitudinal questions deployed through quantitative methods such as the General Social Survey (GSS)[3] (Glaeser et al, 2000 [53]). Results from investigations that have used this approach tend to indicate that significant variance exists in propensity to trust across cultural background, personality type and levels of the social capital that are present within a given community (Hofstede, 1980 [62]). Typically, the source of a persons initial propensity to trust can be decomposed into the following influences: Parents / principal carer, personality type, cultural and socio-economic background:

Parents / principal carer: Research by Mischel (1961) [103] and Mahrer (1956) [87]

[2] *Propensity to trust* is also often referred to synonymously as *disposition to trust*.

[3] Within the GSS, a snapshot of trust is typically obtained using a scaled response to the question: *generally speaking, would you say that most people can be trusted or that you can't be too careful in dealing with people?*

found that young people who had experienced higher degrees of fulfilled promises from parents / authority figures throughout their life showed higher degrees of generalised trust than those who did not.

Personality type: Research by Hofstede (1980) [62] suggested that an individual's propensity to trust is based on two core personality characteristics: optimism, and the perceived capacity to control their own life.

Cultural / socio-economic background: Cultural factors are a known contributor to general trusting attitudes. North American and Japanese citizens have been found to be more generally / readily trusting than Chinese and French (Egger, 2001 [41], Jarvenpaa & Tractinsky, 2000 [67], Fukuyama, 1995 [49]). Irrespective of income, African Americans have been found as being far less trusting than any other ethnic group (Patterson, 1999 [111], Marschall & Stolle, 2004[90]). Rotter (1967) [124] found religious people to be more generally trusting, and found higher trust in higher socio-economic classes than lower (see also Patterson, 1999 [111], Delhey & Newton, 2003 [32]).

2.4.2 The development of trust-based relationships in specific circumstances: *Situational trust*

Situational trust refers to specific instances where a trust-based relationship develops between trustor and trustee in response to a particular set of circumstances. How situation-specific trust relationships are reasoned upon and come to be formed remains the source of continued debate. Traditionally, economists and game theorists, including Axelrod (1984) [8] considered that instances of trusting behaviour that occur in response to specific situations are reasoned upon using a model of purely calculative rationality. Calculative-based trust theory (e.g. Dasgupta, 1988 [30], Malhotra, 2004 [88], Axelrod, 1984 [8]) thus works on the principle that humans make trust decisions based only upon a rationally derived cost/benefit analysis. Dasgupta (1988) [30] illustrated the logic of this view in his assertion that: "*If the incentives are right, even a trustworthy person can be relied upon to be untrustworthy*".

Models of trust investment behaviour that follow the calculative rationality worldview tend to consider that new trustor-trustee relationships begin with a mindset of guarded suspicion. Following such models, trust-based interactions between strangers are initially limited to a form of trust that is *deterrence based*[4]. In a deterrence-based trust relationship, risks relating to a trustee's potential intentions are offset through the utilisation of institutional control structures that serve to deter the trustee from behaving opportunistically. An example of an institutional control structure would be the legal system, where the deterrent is the implicit threat of legal action against the trustee in the event of their transgression. As successful interactions between trustor and trustee repeat however, the need for such control structures becomes more relaxed. At this point, the trust-relationship can develop to a new stage that is described as being *knowledge based*[5]. In a knowledge-based trust relationship, the need for control structures is thought to diminish as, to the trustor, the likely behaviour of the

[4]Some researchers, including Brenkert (1998) [17] use the term *guarded* to describe deterrence-based trust.

[5]Brenkert (1998) [17]: *Extended trust.*

trustee is perceived as being more predictable based upon past experience. Finally, following these models to their conclusion, as successful interactions increase in number yet still, a further stage of trust can be achieved: *shared-identification based* trust. At this highest of levels, trust is thought to generalise between the parties across different contexts, with the need for external control structures essentially eliminated.

However, despite often clear opportunities to behave opportunistically, people often behave in ways that are not in their own interest. It is thought that this may be as result of evolved emotional predispositions (Akerlof, 1983 [6], Frank, 1987 [48]). As a form of emotional psychological contract (Rousseau, 1995 [125]), the act of investing trust is understood to invoke positive emotional reactions in both trustee and trustor. At a basic personal level, it is considered psychologically rewarding to be considered worthy of the trust of another, as well as it is to consider another person worthy of trust (e.g. Good, 2000 [54], Pettit, 1995 [113]). Conversely, when investments of trust are violated, there is often a strong negative emotional reaction; a reaction sometimes strong enough to warrant the complete dissolution of a previously well established trustor-trustee relationship (Sheppard & Sherman, 1998 [135]). The contribution of affective reasoning to trusting attitudes is something that recent models of trust development have now become increasingly aware (e.g. McAllister, 1995 [96]). Social scientists including Lewis & Weigert (1985) [84], and an increasing number of economists (e.g. Nooteboom, 2005[110]) now consider trust to be more complex than the traditional economics models suggested, with both a calculative 'cognitive' component, based upon rational reasoning and an 'affective' component that is based upon a strong positive affect for the object of trust. "*Trust in everyday life is a mix of feeling and rational thinking*" (Lewis & Weigert, 1985 [84], see also Zajonc, 1980 [154]).

2.4.3 The development of spontaneous trust relationships with no prior interaction experience: *Initial situational trust*

Initial situational trust refers to the phenomenon of spontaneous high trust-based interactions occurring person-person or person-institution where there is no prior interaction history between the parties involved. Given the novel nature of pervasive computing, it is this manifestation of trust to which this thesis is primarily concerned. Until relatively recently, and as discussed in the previous section, investments of trust that occurred spontaneously in response to a specific situation and context had been considered as being rooted in a larger developmental process. Within this worldview, high levels of trust (i.e. trusted investments that are made in scenarios of high risk) could only be achieved as repeated successful interactions reduced the perceived risk and uncertainty involved with dealing with a particular trustee (Lewicki & Bunker, 1995 [83], Brenkert, 1998 [17]). There are however instances where high levels of trust can occur spontaneously between parties with no prior interaction history. This phenomenon is referred to as *initial situational trust*. Research into initial situational trust originates from game theory, including work by Berg et al (1995) [14], Mayerson et al (1996) [92] and Pillutla et al (2003) [115].

Empirical studies of initial situational trust behaviour originated from explorations of game theory (Morgenstern & Von Neumann, 1953 [107]), specifically with regard to the *investment game* paradigm (Berg et al, 1995 [14]). The *investment game*, as described by Berg et al

(1995) [14] is a single-shot exercise involving two anonymous participants who have no prior experience of one another. Though several variations exist, the game is typically played thus:

1. Two players / participants are both given some sum of real money (e.g. £10), ostensibly as a 'showing-up' fee.
2. Based on the understanding that each pound sent would be tripled by the time it reaches participant B, participant A is asked to decide how much (if any) of their £10 they would like to invest to the (anonymous) participant B.
3. Upon receipt of whatever investment is made by A, participant B is then asked to decide how much of the (tripled) money received to keep, and how much to send back to A.

Rational logic would suggest that the incentive to reciprocate on any investment made by A would be low as future interactions between the same two parties were not expected, and no explicit deterrent from opportunism had been made explicit. In such a scenario, the rational choice would be for the trustee (participant B) not to reciprocate on any investment from the trustor (participant A), as to perform such an action would only incur an unnecessary loss. Further, a purely rational trustor (participant A) would already have anticipated this and would not make any initial investment to begin with. However, what Berg et al (1995) [14] instead found is that cooperation between people playing the investment game does occur, and occurs often; only two of the 32 participants involved in their study sent no money whatsoever, and many sent their entire stake. Moreover, there was a strong correlation between the amount initially invested and the amount sent back; the higher the sum of money invested, the higher the sum of money reciprocated. People, it would appear, like to trust, and those to whom trust is invested take at least enough pleasure from that investment to provide some degree of reward.

In response to this phenomenon, Mcknight et al (1998) [101] developed a theoretical model of how initial-situational trust might operate. Though their model was originally designed with respect of organisational relationship development (1998) [101], it has subsequently been applied to computer mediated communications, in particular e-commerce (2002a [98], 2002b [99]). The McKnight model considers initial-situational trusting behaviours as being formed of, and influenced by characteristics of both the trustor and the trustee. Within their model, the likelihood of an initial situational trust-based relationship forming spontaneously is informed by the trustor's assessment of four sources of information:

1. The trustor's propensity to trust people / institutions generally.
2. Prior experience either with the specific trustee *or similar*.
3. The presence and perceived effectiveness of usable risk-mitigation devices (e.g. institutional safeguards).
4. The identification and assessment of available evidence as to the intentions and competence of the trustee.

2.5 Facilitating initial-situational trust in computer mediated communication

In all instances of initial-situational trust, be they human-human or human-computer, the trustor must reach a decision quickly as to the degree to which they consider the trustee as being worthy of their trust. This decision is informed in part by the trustor's generalised trusting attitude (Lee and Turban (2001) [82]). Through the use of participant self-reported levels of propensity to trust, some researchers, including Egger (2001) [41], Kim et al (2004) [71], Kim et al (2009) [81] and Mcknight (2004) [102] have found that general trusting attitudes are indeed a useful indicator of initial-trust formation with regard to technologies such as e-banking and e-commerce. However, taken alone, a person's propensity to trust is not generally held to be an accurate determinant of whether trust will be invested in every given circumstance (Johnson-George and Swap, 1982 [69]). Further to propensity to trust, following the McKnight and Chervany model, a substantial contribution to the development of initial situational trust is facilitated by the reasoned (both cognitively- and emotionally-based) evaluation of a number of evidential 'cues'. Such cues are gleaned both from the trustee, and from the situation in which the interaction occurs.

In human-human interaction, the human face is perhaps the most important source of evidence as to the underlying intentions of a potential trustee. Darwinian accounts of the evolution of human emotion propose that the exhibition of inner emotional states are made manifest (and are thus externally detectable by others) through a range of involuntary / unconscious behaviours. Such behaviours include pupil dilation and facial flushing. Being difficult for a trustee to mimic or control, a trustor's ability to detect such behaviours can provide a useful mechanism for the trustor to make an accurate interpretation of the underlying intentions of the trustee – importantly including whether they intend to lie, cheat or deceive. As result of this, the observation of facially based cues is well known to have a strong influence on the degree to which the owner of that face is likely to be considered as being trustworthy.

Reaction to facially based evidential cues is thought to occur rapidly, and at the affective rather than cognitive level[6]. The logic behind the existence of what is essentially a natural lie-detection device is straightforward. As actions based on the misinterpretation of evidence of deceptive intent could have potentially lethal consequences, the ability to accurately detect potential 'cheats' would confer a useful survival advantage to people living within groups. It is thus not surprising to find a degree of support for the existence of evolved cognitive mechanisms designed to detect cheating within Biology and evolutionary Psychology (e.g. Barkow, Cosmides and Tooby, 1992 [10], Macy & Skvoretz, 1998 [86], Boone and Buck, 2003 [15], Debruine, 2002 [31]).

[6]The precise mechanisms involved in the detection of physical exhibitions of behavioural intent are not yet fully understood. Recent advances in functional neurological imaging such as functional magnetic resonance imaging (fMRI) have recently identified the *amygdala*, a structure located within the medial temporal lobe of the brain (and previously known for its role in emotional learning and memory) as being involved in a range of social judgements – primarily those based on facial stimuli (Phelps et al, 2006 [114]). There is now evidence to suggest that the amygdala may be involved in some measure of the perceptual processing of facial features during a trustworthiness evaluation based on facial cues (Adolphs 2002 [5], Winston et al, 2002 [152]).

The communication of cues that are useful in the evaluation of trustee intention (and competence) have also been found to occur in several verbally based behaviours. Cassell & Bickmore (2000) [20], investigating communication behaviours that are involved in human interpersonal trust development, highlight a number of common verbal communication rituals as being involved in assessments of trustworthiness. Such communication rituals include 'small talk', the seeking of 'common ground', and efforts to invoke a sense of credibility through (for example) the appropriate use of technical jargon. All of these strategies share common goals of developing a sense of *rapport*, and also key into a human tendency to categorize and 'unit-group'. *Rapport* is a sense of commonality of perspective, of being harmoniously 'in-sync'. People within a state of rapport tend to consider one another as being in-group and thus tend to look more favourably upon one another than when rapport is absent. Consequently, it is unsurprising that the active development of rapport is one technique to which salespeople rely heavily when attempting to pitch their wares to potential customers. *Unit grouping* (McKnight and Chervany, 1998 [101]) refers to how people within specific groups tend to share similar values and goals, and how this leads individuals to be more likely to look upon in-group members more favourably than out-group members (Brewer and Silver, 1978 [18]). Unit-grouping in particular is well known to be an important means by which people assess the trustworthiness of other people: *if they are like me, and I consider myself trustworthy, then they are more likely to be trustworthy* (McKnight and Chervany, 1998 [101]).

2.5.1 Initial situational trust formation within interactions that are conducted in different places and times

With respect to facilitating the greatest range of potential evidence to a trustee's trustworthiness, negotiations of trust might well be thought of as being best conducted face-to-face, in the same place at the same time. In many situations however, performing a negotiation of trust face-to-face is simply not possible. In such scenarios, people must find their evidence for trustworthiness by other means.

Computer mediated communication (CMC) allows people to perform asynchronous transactions across vast distances of time and space. While this offers considerable benefits, not least in terms of convenience, it also places limits upon the ways people are able to establish trustworthiness using the techniques so far discussed. Giddens (1990) [51] refers to the mediation of transactions across time and space in terms of degrees of *embeddedness*. As the level of mediation between the parties that are involved in an asynchronous exchange increases, the interaction is considered as being increasingly *disembedded*. As a transaction becomes increasingly *disembbeded*, the demands that are placed upon trust will be increased (Riegelsberger & Sasse, 2001 [117]), with a corresponding additional demand on the evidence that would be required to support it. In the following sections, a number of methods by which evidence of trustworthiness can be conveyed to users of CMC are presented and discussed.

Virtually re-embedding the interaction through the modulation of facial cues into CMC

Until relatively recently, the technical limitations of computer mediated communication meant that many of the more subtle cues that people use to establish trustworthiness (facial cues being the most obvious example) were severely impoverished or simply absent. However, as the range of communication modalities afforded to systems such as the WWW have increased, designers of user interfaces and interactive technologies have been able to use modulated versions of human interpersonal trust cues within their systems. The modulation of interpersonal human-human trust cues interaction to the (more dis-embedded) human-computer domain has been described by Riegelsberger & Sasse (2001) [117] as a process of 'virtual re-embedding'. Virtual re-embedding has been discussed as being achievable in two general ways: 1) through improving communication transparency, and 2) through the imbuing of personal human trust cues to virtual agents though anthropomorphism.

Systems such as the WWW can support both methods of virtual re-embedding in a number of ways. Being graphically based, websites can present digital representations of people, either as static photos or as synthetic representations of humans that can then be anthropomorphised). The degree to which the inclusion of such imagery can increase perceptions of trustworthiness has been examined by a number of researchers in a variety of web-based contexts. Within interface design for systems heavily reliant upon trust (e.g. online banking, e-commerce), Fogg (2001 [46]), Zheng et al (2002) [157] and Steinbruck et al (2005) [142] all found that the inclusion of photos of 'staff' incurred a positive affective reaction in participants that viewed them. Inclusion of such imagery was later found to increase user perceptions of the trustworthiness of the company whose staff the images were purported to be representing. However, results have not been entirely conclusive: Riegelsberger et al (2002) [118] found no general significant effect upon perceptions of trustworthiness, and indeed some evidence that the inclusion of staff photographs had a negative effect upon trustworthiness. Similarly, while Hertzum et al (2002) [61] found strong positive affective reactions to virtual agents who were represented by photos of real people, negative reactions were found to be associated to those that were represented by an artificial personification (e.g. computer-modelled and animated).

Indicators of trustee investment

Through the act of seeking to engage in a new trust-relationship, the sense that a trustee is (directly or indirectly) investing something they would not wish to lose is a powerful means of creating a perception that they are more likely to be trustworthy. In CMC, two common indicators of trust investment are *organisation size* and *reputation*.

Organisation size: In the commercial world, there is a strong positive correlation between the perceived size of a company and the degree to which people consider that company to be trustworthy (Doney & Cannon, 1997 [39])[7]. Large commercial companies take time (and

[7]Egger (2001) [41] has also suggested that the legal status (e.g. corporation or public limited company etc), the associations that company holds and / or the endorsements of high profile customers are all similarly effective in this regard.

resources) to build and develop, and their continued existence is often largely dependent upon the sustained goodwill of its customers. Perceptions of large company size has been found to be positively correlated to subsequent intention to trust by Koufaris & Hampton-Sosa (2004) [79], whose participants noted that size conveys a sense of competence, as it creates an impression that they possess the capabilities necessary to provide the products and services expected of them. Further, a company of large size was perceived as being in a position to offer increased levels of structural assurance, such as compensation in the event of failure to fulfill the service that they offer.

Reputation: Reputation is that human trait that helps to "*manage the complexity of social life by singling out trustworthy people, in whose interests it is to meet promises*" (Misztal, 1998 [104]). As one's reputation is a trait that takes time and energy to develop, it is generally perceived by a trustor to be something that a trustee would consider as worth protecting. Reputation-based trustworthiness effects exist because, as Axelrod (1984) [8] describes, there then exists a '*shadow of the future*' – i.e. the prospect of future retaliation or loss of reputation as result of behaving opportunistically. In the commercial world, a companies reputation is often deliberately conveyed to customers through use of a brand identity (Koufaris & Hampton Sosa, 2004 [79] Chen & Barnes, 2007 [24]), upon which considerable time and money will likely have been invested.

In WWW based scenarios such e-commerce and e-banking, conveying evidence of investment has been found to be important to customer initial trust. Researchers, such as Fogg [47] have examined the use of such cues extensively, and a summary of the more salient findings are presented thus:

Aesthetics: Investigations by Mcknight (2002) [99] and Koufaris & Hampton Sosa (2004) [79] have both found a positive relationship between site aesthetic quality / appeal and reported intentions to trust in initial-situational trust-based scenarios involving e-commerce. The specific mechanics of aesthetic appeal has been explored extensively by Fogg (e.g. Fogg et al, 2001 [47]), Riegelsberger and many others. A professional looking interface is thought to create a positive affective reaction as it implicitly conveys to the user that time (and most likely money) must have been invested by the individual or organisation which that website represents. Though the specific mechanics of what constitutes 'professionalism' are rarely detailed, adherence to design principles such as the correct use of layout and space, alignment, typography, consistency and the quality of imagery used (and by extension the likely cost involved in sourcing it) could all form evidence that design expertise has been employed (and, by extension, paid for). Conversely, inconsistent visual design and technical failures are both highlighted by Riegelsberger & Sasse (2001) [117] as being a source of strong negative affective reactions, with a corresponding reduction in reported initial-trust intention attitudes.

Functionality: The usefulness or benefits provided through use of the service have consistently been found to positively affect initial trust (Kim et al, 2009 [81], Koufaris & Hampton Sosa, 2004 [79], Chen & Barnes, 2007 [24]), as has the inclusion of functionality that allows users to customize or tailor products / services to individual taste (Koufaris & Hampton Sosa, 2004 [79], Chen & Barnes, 2007 [24]).

Usability: Though the degree to which the usability of an interactive system is related

to initial trust is not clear (Koufaris & Hampton Sosa (2005) [59] for example found no link between initial trust and usability), failure to consider the ease-of-use of an interactive system has been highlighted by Riegelsberger & Sasse (2001) [117] as forming a potential '*trustbuster*' with respect of initial-situational trust.

Provision of structural assurance: Structural assurances offer routes of risk mitigation such as auditing, contracts and insurances. The presence of such assurances have all been found as being very important in initial trust (Kim et al, 2009 [81], Riegelsberger, 2005 [120] Mcknight et al, 2002 [99], Kim et al, 2004 [71]), and particularly so with the relatively high risk services of e-commerce and e-banking. However, how such assurances are conveyed to users is important. Riegelsberger (2005) [120] found that the inclusion of a salient personal trust cues without associated trust functionality is poorly received by users, and in a study of e-commerce websites conducted by Mcknight (2004) [102], neither a noticeable TRUST-e privacy seal nor a noticeable professional association seal were found as having any significant impact on initial trust in the websites tested.

Indicators of trustee status

In the socio-physical world, the presence of clear identifiers of social status, such as a uniform, ID card or rank insignia can serve as a highly useful visual indicator of trustworthiness in that they convey a sense of authority and, by extension, potentially other positive character traits such as integrity. Within computer-mediated communication scenarios such as the WWW, *trust seals* are a popular modulation of a salient status identifier. *Trust seals* are a visual cue that are included in a website to provide evidence that that website has been checked and validated by an independent institution (two examples are *Verisign* and *Trust-e*). By following the link that the seal provides, users of the website can further validate this information by checking corresponding information on the independent institution's website.

Status and authority key into a common heuristic sometimes referred to as the 'white coat effect': i.e. if a person in authority says it is OK, then it is probably OK. However, whilst powerful, the use of status identifiers as cue to trustworthiness remain common cause of mistrust errors generated from their over-reliance. The ease with which graphically based design elements on the WWW can be faked, and the difficulty in attracting user attention to browser based cues (that are of course outside the main focus of the user's attention) remains a continuing source of mistrust errors and a major contributing mechanism to confidence-trickery based technology attacks such as 'phishing[8]'.

Reducing uncertainty through increasing a trustor's control over the interaction

A significant contribution to a trustor's reasoning about trust, particularly initial-situational trust is the degree to which they perceive the outcome of co-dependent action as being uncertain. Earlier in this chapter, the core source of uncertainty as it relates to trust was

[8] *Phishing* describes an attack whereby a masquerade of an electronic communication, purporting to originate from a trustworthy person or institution, is used to fraudulently capture private / personal information through trickery. The use of the phishing attack is discussed in detail in chapter 6.

discussed as being directly attributed to the inability of the trustor to attain complete control over the trustee and the interaction. Perceptions of risk are generally higher in situations of low control (Corritore et al, 2003 [27]), thus by increasing the perceived level of control, the perception of risk involved in the interaction should be reduced (Egger, 2001 [41]). In human interpersonal trust, the term *token control efforts* describe a type of micro behaviour that can used to probe a potential trustee to see how they react to the influence of the trustor (McKnight & Chervany, 1998 [101]). Such behaviours are useful to trust as they are able to covertly 'test the waters' of future interactions without the need for investment or the assumption of any actual risk. The objective of token control efforts is to create, in the mind of the trustor, a perception of the degree to which the trustee is likely to be predictable and, in doing so, support an internal locus of control (Shapiro, Sheppard & Cheraskin, 1992 [132]). Such probes are often extremely subtle. An example is attempting to make the trustee reciprocate on a simple gesture, such as a laugh or smile (McKnight & Chervany, 1998 [101]).

As Rutter (2001) [126] notes with regard to trust and technology, it is important to maintain a sense of 'knowing what is going on' in the mind of users who are entering a situation where risk is present. With regards to technology, this is especially true during times when a system makes demands of a user without offering anything upfront or in return. Benefits associated with the provision of predictability and the supporting of a user's internal locus of control is a common occurrence in the user-interface design literature (e.g. Schneiderman's '8 golden rules' [136]) – particularly with regard to e-commerce website design. In such interfaces, typical recommendations include the inclusion of salient escape routes (allowing users a clear way out of a process at any time, without penalty), improved transparency of link action (so users are aware of the action that will be performed when a link is clicked), the permission of 'dry runs' and adequate feedback to actions that are performed by the user (Schneiderman, 1997 [136], Riegelsberger, 2001 [117]).

2.6 Initial situational trust in pervasive situated services

A number of large scale research projects have examined pervasive computing-based infrastructure and service usage across a range of scenarios and potential uses. Some of the research conducted within these projects explicitly examined the role of human trust with regard to pervasive computing service usage.

Equator: As a six year long interdisciplinary research collaboration, the *Equator* project (2000-2006) examined a wealth of issues relating to ways in which the physical and digital worlds could be integrated. With a broad remit and with over 60 researchers actively engaged, *Equator* examined many uses of pervasive technology, including for learning / education (e.g. Rogers & Price, 2006 [123]), gaming / entertainment (e.g. Crabtree et al, 2004 [29], 2007 [28]), assistive technology (e.g. Cheverst et al, 2003 [25]) and computer supported collaborative work (e.g. Rodden et al, 2003 [122]). Several projects conducted within *Equator* involved aspects of interpersonal and human-computer trust, with an excellent example being the mixed-reality game "Uncle Roy All Around You" (Benford et al, 2004

[13]), that involved trust (specifically trust in strangers) as a central theme. In "Uncle Roy All Around You", players were deliberately drawn into situations where their investment of 'well-placed' trust (i.e. in the veracity of information presented to them from a number of different sources) would be crucial to their progress in the game.

Cooltown: The Hewlett Packard Labs *Cooltown* project (e.g. Kindberg et al, 2002 [72], Barton & Kindberg, 2001 [11], Spasojevic & Kindberg, 2001 [140]) developed a pervasive computing infrastructure that was able to offer visitors to a purpose built laboratory a rich interaction experience based around what was described as a *real-world wide web*. Within *Cooltown*, the laboratory structure (walls etc), and a range of physical objects within the laboratory were afforded web-presence by means of embedding web-server capabilities within them. In doing so, objects within the physical world were afforded a link to corresponding information pertaining to them in the digital world. Using this infrastructure, researchers involved in the *Cooltown* project were afforded a large scale pervasive computing environment within which they could develop and test a range of new services and techniques. Output from the *Cooltown* project included a user evaluation of novel wireless payment mechanisms (Kindberg et al, 2004 [75]), to which user trust was found to be a strong influencing factor. Similarly, **Mobile Bristol** (also involving Hewlett Packard Labs) examined a number of different ways that mobile technologies could be used for enhanced interaction with the physical environment, using mobile devices in conjunction with a range of deeply embedded sensor technologies.

However, though some of these projects did seek to tackle the role of human trust with regard to pervasive computing service usage to varying degrees, only very few (e.g. "Uncle Roy All Around You", Benford et al, 2004 [13]) directly examined the role of initial situational trust (though not explicitly). The characteristics of pervasive computing systems generate a number of new challenges for initial situational trust development that have not been reflected in the techniques thus far discussed. Two of these characteristics are especially important:

Digital service discovery: In a pervasive computing world, the presence and availability of particular digital services in a given physical space may not be made immediately apparent. The user may first need to 'discover' proximate digital services before they can be used.

Digital service authentication: Before a digital pervasive service can be used safely and securely, it must be authenticated by the user as being genuine, or at the very least, verified as being the service to which the user had intended to connect. In an ideal situation, this process would be performed before a full connection is established between the user's device and the host's server.

2.6.1 Digital service discovery

Not unlike the WWW, within a pervasive computing environment a user will, within a given time and space, come to encounter a number of co-existing services from which they can choose. Less like the WWW however, the range of pervasive services that may be available in any given physical space (and perhaps time) is far more dynamic. Pervasive services are

dependent on an ever changing population of devices and the ad-hoc infrastructure they are able to provide. However, how pervasive services will come to be discovered and interacted with must ultimately be handled by the user's device via some model of service discovery.

Protocols that handle the discovery of pervasive computing services will need to perform many roles, including the efficient management of data traffic within an environment where the availability of communication infrastructure is constantly changing, as well as the protection of devices, services and users (Zhu et al, 2005 [158]). At the technical level, the development of protocols to better support pervasive service discovery is subject to continued and sustained research effort. Though protocols that offer service discovery functionality appropriate to the pervasive domain have been offered (e.g. Chakraborty et al, 2006 [23], see Zhu et al, 2005 [158] for a review), it will be some time before they find large-scale commercial use.

2.6.2 Spontaneous secure ad-hoc device association

Pervasive systems, in their purest sense, are 'everywhere'. Oftentimes, their discovery and usage is conducted via connections that are made wirelessly. However, the act of removing the physical cable that serves to 'tie' communicating devices together, while practically convenient, is also the source of a new set of problems.

In "Security and trust in mobile interactions", Kindberg et al (2004) [75] explored user perceptions of trust in an experimental setting using a simulated cafe as part of their research strategy. In their study, a number of different ways of allowing 'customers' to make payments were evaluated. Several of the methods involved using wired connections, while several others used wireless connections. Analysis of participant reasoning indicated a concern with wireless connections; participants were concerned that they might feasibly connect to the wrong service. In this regard, wired services were often considered more secure. They also found that the relationship between user perception of risk and actual risk was somewhat loose; understanding and awareness was typically poor, and participants tended to trust what was visible, tangible & apparently human-free.

As Kindberg (2003) [76] notes, the lack of a physical link between devices that are communicating on a wireless network removes any immediate physical indication as to which device is at the other end of a given connection. In a pervasive computing scenario, this problem is compounded somewhat further by the possibility that many independent digital services and many mobile devices can be co-present in any given space. How then, can the user of such services be sure that they are indeed connected to the service / device to which they had intended to connect?

This particular use-case scenario, the creation of user-verifiable ad-hoc device associations (e.g Kindberg and Zhang, 2003 [76]) was initially discussed by Stajano & Anderson (1999) [141]. The solution to this problem is far from trivial. From the user's perspective (in terms of the degree to which they trust the connection), the method by which connections are made must be both efficient and intuitive yet, from a technical standpoint, solutions must also be secure. Catering for both is made difficult because the technical requirements

that underpin the security of digital communications are often complex. For spontaneous device associations to be considered secure, they must be able to resist two general types of threat: *The man-in-the-middle* (MITM) and the *evil-twin* (ET). The *man in the middle* attack refers to communication eavesdropping, whereby communications from person 'A' are intercepted by the attacker en-route to the service 'B'. Acting as the the *man-in-the-middle*, the attacker remains - undetected - as a relay, collecting data traveling both ways, before relaying that data on to its intended destination. Conversely, the *evil-twin* attack describes the creation of an entirely bogus digital service that masquerades as a similar bona-fide service. An attacker can perform the *evil-twin* attack in two general ways:

1. The attacker creates a shared wireless internet gateway service that spoofs / mirrors the bona-fide service's identity by copying its name[9] exactly. He then attempts to overwhelm the radio signal of the bona-fide service, perhaps through use of a more powerful signal generator.
2. The attacker simply creates a competing service with a similar name / SSID to the bona-fide service (e.g. *Starbucks WiFi* as opposed to the genuine *Starbucks Wireless*).

Both types of threat described have significant security implications attached to them. In either scenario, as the attacker is able to offer a service with the same characteristics as the bona-fide service (perhaps even including the same data encryption services), it is difficult for a typical user to notice anything that is untoward. However, the mere act of connection may be enough for an attacker to compromise the personal / private data of the user and / or install malware on the users device.

User-verifiable device association through use of 'out of band' auxiliary communication channels

Since (and including) Stajano & Anderson (1999) [141], a number of protocols have been developed by researchers that seek to provide means of enabling ad-hoc device associations that are able to combat the threats posed by the MITM and ET attacks[10]. Many such methods directly involve the user in the authentication process directly.

A number of methods for secure spontaneous device association have explored the use of 'out of band' communication channels that enable user verification without the need to expose or compromise the primary communication channel (and thus potentially compromise the data stored on the initiating device). Typically, the ultimate function of the 'out of band' channel is to swap a small amount of information between devices in a way that the human user can verify that it is the intended target device that is being communicated with. Upon verification of this initial communication by the user, the information that was exchanged using the 'out of band' channel is subsequently used to create a full secure device pairing using the primary communication channel.

[9] The 'name' of a wireless service is provided as its SSID (Service Set IDentifier), which can be set by the service administrator.

[10] A review of such techniques has recently been offered by Kumar et al (2009) [80] and a comparative user study was conducted on a number of proposed pairing methods by Kobsa et al (2009) [78].

Using the auxiliary channel approach, the side channel is first used to perform the initial connection between the user's device (the initiator) and the target device. Upon making this initial connection, the role of the user is then to verify that a connection has been performed with the intended target device. To do this, the auxiliary channel must offer some form of human-perceptible evidence[11]. Two general approaches have been taken in respect of the use of auxiliary channels as a means of conveying evidence of device association. In the first class of approach, the auxiliary channel itself forms the evidence by limiting the ability for an attacker to intervene in the interaction through using channels that are location-limited or otherwise physically-constrained in some way.

Location limited channels: Location-limited channels (e.g. Balfanz et al, 2002 [139]) are deliberately short-range communication channels. As such, they rely on the user's intuitive belief that they are better able to eliminate the *man-in-the-middle* type attacker by simply offering little in the way of physical room within the interaction for them to intervene. Several solutions, including Stajano & Anderson's (1999) [141] 'Resurrecting Duckling' facilitate device pairing by simply placing the devices in direct physical contact with one another. Similarly, Riekki et al (2006) [121] and Balfanz et al (2002) [139] have used near physical contact, using short range communication technologies such as Radio Frequency Identification (RFID) tags. As a final example, McCune, Perrig & Reiter (2005) [97] explored the use of 2D barcodes. In their protocol, the digital camera facilities of the initiating device (a mobile phone) were used to pair with a host by reading a 2D barcode that was displayed on the screen of the host device (also a phone) - again severely limiting the ability of an attacker to intervene.

Other variations of location-limited device pairing have extended the security offered by direct device contact by further utilising information about some shared characteristic of the devices involved in the negotiation. Mayrhofer & Gellersen (2007) [93] expanded a direct contact protocol by combining the contact with an elegant key generation technique that used accelerometers embedded in both communicating devices. In their protocol, communication is first established by placing the devices together, and a unique cryptographic key is then generated through the vigorous, random shaking of the devices in tandem (see also 'Smart-ITs', Holmquist et al, 2001 [63]). This key is then used to secure subsequent communications between the devices. Other solutions by Mayrhofer, Gellersen & Hazas (2007) [94] and Kindberg & Zhang (2003) [77] have used measurements of the relative spatial location of each device, derived using ultrasound. As a final recent example, Varshavsky et al's (2009) [148] "Amigo" protocol used an even more sophisticated technique whereby communicating devices would derive and later compare a shared radio fingerprint based on the immediate characteristics of their surrounding radio landscape.

Physically constrained channels: In situations where direct, or near direct contact is either impractical or unavailable, several solutions have also been posited that do not require communicating devices to be immediately proximate. Typically, such solutions instead utilise communication channels that are physically-constrained in some way. Some physically constrained channels that have been explored include laser (Mayrhofer & Welch (2007) [95] and Kindberg & Zhang (2003) [76]) and infra-red light (Balfanz et al (2004) [9]). As the communications conducted over the auxiliary channel are insecure, it is the precision of

[11] Balfanz (2002) [139] has described this process as the '*demonstrative identification of communicating devices*'.

Figure 2.1: *Visualising hash keys for easy comparison*: In the examples shown, an original Md5 hashed key is compared with two others (one identical, one not). By modulating the raw key to an image, identifying the odd-one-out is made much easier for the human perceptual system to detect.

the physically constrained channels that is the source of its evidential value. In both the laser and IR examples mentioned, communication between initiator and host devices must therefore be conducted line-of-sight.

In the second class of approach, the auxiliary channel is instead used to facilitate user authentication by presenting evidence that the user is able to compare between the initiating and host devices, so as to establish their direct association with one another. Underpinning this approach is the initial creation of a cryptographic key that is shared between the initiating and host devices. The keys that are held by each device are then presented back to the user for comparison using the auxiliary channel.

Various modalities by which such verification can be achieved by direct comparison have been explored, including the use of various forms of visual (see figure 2.1), auditory and tactile feedback. Through modulation into an image for example (e.g. Perrig & Song, 1999 [112]), 'visual hashes' have been found to be highly effective as a means of comparison as compared to their raw string state (e.g. Dhamija, 2000 [36]).

2.7 Chapter summary

In this chapter, a review of literature pertinent to human-computer trust was conducted, providing an initial means of addressing **RQ1**: *How is human-human interpersonal trust understood to operate, and how has this been modulated for use in human-computing applications where some degree of trust (on the part of the human) is obligated?* Through this investigation, the output of a number of research streams that have sought to understand human trust and the ways in which cues that support it can be modulated for use in technologically-mediated communication were presented and discussed.

Trust is a form of rationality that allows people to perform actions in situations of risk and

uncertainty. That a person comes to invest their trust in someone or something in a given situation is driven by a subjective measure of *trustworthiness* that is ascribed to a trustee from the trustor. The value of trustworthiness that is ascribed to a trustee is derived from a combination of cognitive and affective reasoning about both the trustee and the situation within which trust is to be invested.

Trust-based relationships typically develop over time and repeated interactions, with high degrees of interpersonal trust usually only occurring as repeated successful interactions reduce the degrees of risk and uncertainty that are inherent in such situations. However, the phenomenon of *initial-situational trust*, in which high degrees of trust are found to occur spontaneously in situations where the trustor has no prior experience with the trustee, can occur in certain circumstances, providing certain criteria are met.

That interactive systems are considered as being trustworthy is an important factor in how users come to accept and use those systems. This is particularly important with regard to systems where the privacy and security of a user's data might be compromised through their use. With the advent and continued advances of the services that are now available on technologies such as the WWW, various methods by which trustworthiness can be imbued into the designs of risk-relevant systems have been explored. Indeed, by utilising the full range of communication modalities that are now afforded to the WWW, systems designers are able to modulate several well known human-human trustworthiness cues into the human-computer domain with positive effect.

However, several new characteristics of pervasive computing mean that services deployed using such infrastructure are currently much less capable of offering the same level of immediate evidence as to their trustworthiness as might be found on systems deployed via the traditional desktop-metaphor. Such characteristics, include the need for users to both discover and authenticate the services that the pervasive computing infrastructure of a specific place and time can offer. Though many novel methods have been presented to address trust issues related to these new characteristics, there is currently little research into how users might consider them in terms of their trust.

In the next chapter, the focus of the investigation will turn toward gaining an insight into user attitudes and behaviours regarding current uses of technology within situations where the use of trust is particularly salient.

Chapter 3

An examination of current user experience and behaviour with technologies that involve issues of personal privacy and security

3.1 Chapter overview

In terms of their impact into the public consciousness, pervasive computing services, such as public access WiFi are a relatively new phenomenon. As such, gaining insight into how people perceive them, particularly in terms of the issues of privacy and security that are associated with them is difficult. In this chapter, an initial investigation sought to uncover the salient privacy and security concerns relating to traditional desktop-metaphor computing that would likely transfer into similar types of activity conducted in a pervasive computing scenario.

In this chapter, the researcher conducted an investigation to begin developing a deeper understanding of the degree to which user behaviour with familiar ICT technologies might contribute to the development of initial-situational trusting attitudes with regard to more nascent pervasive technologies such as *situated services*. To aid the study, a questionnaire was developed to gather information about the degree to which people currently use interactive ICT (such as the WWW), and the degree to which they understood and considered threats to the privacy and security of their personal data as a result of their technology-related behaviours. Further, the questionnaire also sought to gather information as to the measures that users currently took to protect their data as a result of their technology-related activities. In terms of the thesis as a whole, the results of this investigation will address **RQ2**: *To what extent do people currently understand the threats and risks associated with*

the use of ICT services such as the WWW? To what extent does this understanding translate to pervasive / situated computing services?, and will also be used to inform the design of the empirical component of the research discussed in later chapters.

3.2 Introduction

The private and / or personal data of technology users is a valuable commodity to those who are keen to exploit technology for malevolent purposes. When encountering a new technology, users should always consider the degree to which their personal / private information is secured, and how their use of that technology might affect that security. From the perspective of the system designer, the degree to which users of technology consider the potential risk of loss or misappropriation of their personal information is also important; it will likely be a contributing factor to their decision of whether (or indeed not) to use that technology, and thus directly affect whether that system will be accepted.

Despite the Internet remaining largely open and ultimately unregulated, many activities that involve risk - and particularly those activities that involve financial risk - are now commonly and routinely performed by people using services that are made available on the WWW. The continued usage growth of risk-relevant web services (such as e-commerce and e-banking) alone provides us with evidence that people are, to the most part, comfortable with conducting activities that involve entrusting their personal / private data to others. However, as the ability to perform risk-relevant activities (such as the payment of goods) moves from the desktop at home and into public space, issues of security and privacy to which many people have become accustomed will need to be re-evaluated. In a pervasive computing-based world for example, users may be obliged to expose valuable personal information on networks to which they have little personal control, possibly using communication devices that are not their own. While the effects of this are not yet well understood, there is evidence to suggest that people consider such a move to be uncomfortable. Nilsson, Adams and Herd (2005) [109] for example have found that people using online banking services were generally only happy to use them in certain locations, usually their home, and not on other peoples / publicly accessible machines. Similarly, Kindberg et al (2004) [75] encountered significant user concern relating to use of wireless connectivity as a means of facilitating customer payments.

The Mcknight & Chervany model of initial-situational trust formation (Mcknight et al (1998) [101]) proposes that the degree to which a trustee (of which the trustor has no previous experience) is considered trustworthy is informed in part by their prior experience with trustees / situations that they consider as being similar. It is this insight to which the investigation discussed in this chapter is primarily based. As ICT services move closer towards pervasive-infrastructure deployment, the role that initial situational trust will play in their early acceptance will likely be critical, initially because of their novelty. Discussing technology use in general, Rutter (2001) [126] found that when people approach a new experience, they often tend to apply rules that have governed similar experiences in a similar domain. Thus, while a person's prior direct experience with pervasive technologies may be essentially non-existent, some of their more familiar characteristics (use of wireless connectivity etc) may prove to be an important factor in how such technologies / services are initially

perceived and their risks understood by users.

3.3 The design of a questionnaire to investigate technology usage behaviour relating to user privacy and security

User acceptance of risk-relevant ICT technologies and services is a highly dynamic process. Particularly with respect to the WWW, though the security measures that support and protect users from malicious behaviour continue to improve, so too do the counter-measures that aim to break or otherwise subvert that security. Nonetheless, as the benefits that the WWW has been able to offer have grown, so too has user acceptance as to the risks of its use. However, a consequence of this highly dynamic developmental process is that it is often difficult for researchers to maintain an accurate picture of current user opinion and practice regarding technology use and acceptance. For example, while the practice of accessing personal financial information through an Internet known to be populated with 'hackers' and other technically superior undesirables might have been considered far too risky only five years ago, it is now widely accepted. Thus, in order to investigate *current* user concerns about privacy and security issues surrounding their current uses of technology, a questionnaire was developed. As an initial investigation into user behaviour related to technology use, an objective of the research presented in this chapter was to gain information from a large number of people across a wide demographic. The standardised and objective nature of questionnaires was considered to be more appropriate than interview at this stage of the research. Deeper explorations of user reasoning relating to privacy and security themes relating to specific instance of technology use will however be explored in later chapters. The design of the questionnaire was conducted using an iterative design process that was based around two core themes. Those themes were: 1) *the extent to which people considered their privacy and security with respect to the technologies that they currently used* and, 2) *the ways in which people feel that systems of a pervasive nature might affect their personal privacy and security.* A set of core questions relating to these themes was first generated using a focus group consisting of the researcher and several research colleagues. The resulting questions were then subsequently refined using several short pilot sessions in which small groups of participants completed the questionnaire and provided feedback as to its comprehension and interpretation.

3.3.1 Activities and behaviour relating to current technology usage that involve issues of privacy and security

In order to capture information about the degree to which people considered their privacy and security with respect to the technologies that they currently used, their actual technology usage behaviour would need to be examined, with particular attention paid to the extent to which technologies that carried a risk to the security and / or privacy of their personal data. To facilitate this, a group of questions were developed to capture information related to four general issues:

- The extent to which people had experience in using technology to perform activities that posed a potential risk to their security and / or privacy.
- Where those activities had taken place, e.g. at home, within public space or both.
- The measures taken by people in order to protect their security and privacy with respect to the technologies that they owned / used.
- The extent to which people had personal experience, or been victims of, technology-based malicious activity (e.g. phishing attacks / credit card fraud etc).

3.3.2 Technology user concerns about pervasive computing services in relation to personal privacy and security

The relative novelty of pervasive services made the choice of questions in this section of the questionnaire difficult. While needing to be reflective of core issues of security and privacy in relation to pervasive computing, they also needed to be familiar in some respect to the the respondents such that they could assess them relative to their current technology-based behaviour. To capture information as to the thoughts and concerns of people relating to technologies of a pervasive nature, a number of technologies were selected by the researcher as being reflective of issues that are pertinent to pervasive computing in general[1]. The issues considered as being most directly pertinent were knowledge of *user location* and knowledge of *user identity*. Two technologies were selected as they related to systems that were, or would be reliant upon, knowledge of the location of people / users (i.e. location-based). These were:

- The use of fixed / mobile closed-circuit television (CCTV) surveillance in U.K. towns and cities.
- UK government proposals for per-usage road pricing, as made possible by the sustained tracking of citizens individual vehicles.

The remaining two technologies were selected as they referred to systems that would rely upon a knowledge of user identity. These were:

- The UK Government's proposal for individual citizen ID cards.
- The UK Government's proposal for biometric data to be included on UK citizen passports.

[1]The technologies that were chosen were current at the time of the questionnaires deployment in the summer of 2007.

3.3.3 Questionnaire deployment

The questionnaire was deployed using an online survey system that was made publicly accessible via the WWW. Invitations to complete the questionnaire were distributed via online University noticeboards and the online social networking website *Facebook*. This questionnaire was self-administered, and thus completed by respondents of their own volition without the researcher present. As several questions were asked in respect of issues relating to the personal privacy and security practices of the respondents, responses to the survey were made anonymously.

As the questionnaire was designed with current users of technology (and particularly the use of services that are made available via the WWW) in mind, deploying the questionnaire within an online environment was considered apt. However, a paper based version of the questionnaire was also made available for completion offline if requested. A copy of the questionnaire as used in the study can be found in appendix A, section A.1.

3.4 Results

Responses to the questionnaire were collected from 229 participants (gender split: M=131 (58%), F=97 (42%)) over a 12 month period between July 2007 and July 2008. Demographic measures were obtained for age, gender and the degree to which respondents considered themselves as being I.T. competent / literate. The modal age range of respondents was 27-36. Self reported measures of perceived levels of personal I.T. literacy (using a 7-point Likert scale: 1 *basic user* to 7 *highly accomplished user*) indicated that 198 (86%) of respondents reported their personal level of I.T literacy as being 5 or above (mode = 7). 83 (36%) of respondents rated their personal I.T. literacy at the highest level (7).

3.4.1 Reports of personal experience of criminal / malicious behaviours through personal technology usage

It is of course important to note that prior user experience with technology can be both positive and negative. Specifically with regard to their use of the WWW, respondents were asked whether they had personal experience of, or indeed fallen victim to three common types of malicious behaviours related to technology usage. These behaviours were: virus / malicious software infection, credit card / other banking fraud and / or a 'phishing' attack[2]. Frequencies were calculated for each of these three types of attack were calculated and are presented as a table in figure 3.1.

All respondents reported as having had personal experience of at least one form of the malicious online behaviours described, and very few respondents stated as being unsure as

[2] *Phishing* describes an attack whereby a masquerade of an electronic communication, purporting to originate from a trustworthy person or institution, is used to fraudulently capture private / personal information through trickery.

Attack type	Yes	No	Don't Know
Phishing attack	172 (75.1%)	48 (21%)	6 (2.6%)
Credit card / banking fraud	41 (17.9%)	182 (79.5%)	6 (2.6%)
Virus / malware infection	132 (57.6%)	93 (40.6%)	4 (1.7%)

Table 3.1: Respondent reports of personal experience of malicious activity through their use of technology

to whether they had fallen victim to any one of the behaviours described or not. 'Phishing' attacks appeared most commonly, with three quarters (172, 75.1%) of respondents indicating that they had, at some point, encountered such an attack. Over half of the respondents (132, 57.6%) reported as having had their technology infected with some form of virus or malware. Reported instances of credit card fraud, though low relative to the other types of behaviours examined, still approached one-in-five (41, 17.9%). However, it was not clear as to whether the fraud was merely attempted, or was actually successful.

3.4.2 Considerations of privacy and security in relation to pervasive technology

The results presented in this section were collated from open-ended responses to questions relating to forms of technology that involved aspects of user privacy / security that were considered as being relevant to themes of pervasive computing in general. One question related to a currently used technology (CCTV), while the remaining three questions were about technologies that were (at the time the questionnaire was available) being proposed by the U.K. Government. These technologies were *per-usage road pricing, as made possible by the sustained tracking of citizens individual vehicles*, *individual citizen ID cards* and the use of *personal biometric data on the U.K. passport*. Though open-ended, responses to the four questions were found to be typically short. As such, the use of a formal methodology to analyse the data generated by each question was not considered appropriate. To present and discuss the responses to each of the questions posed, responses were first assigned to one of four general categories *mostly positive*, *mostly negative*, *mixed* (i.e. containing both positive & negative points) and *ambivalent* (i.e. stated as 'don't care' or similar). Responses in each category were then examined individually, with the most common / frequently occurring themes noted by the researcher as they emerged.

Respondent consideration of privacy / security issues related to technologies that would utilise knowledge about user location

CCTV: In response to the question *How do you feel about the use of CCTV fixed / mobile surveillance in U.K towns and cities?*, roughly a third (48, 28.9%) of respondents considered CCTV in as a broadly positive technology, with a further 35.5% (59) considering it in largely negative terms and the remaining third (59, 35.6%) being either ambivalent or of mixed feelings. Positive aspects of CCTV were most often discussed in terms of its use for

the detection and prevention of crime and anti-social behaviour. The benefits of CCTV as a deterrent from criminal activity was also frequently mentioned, with corresponding comments about general feelings of increased personal safety as a result of this perceived effect.

Respondents who reacted with strong negative opinions about CCTV however frequently mentioned that its presence and use was an invasion of their personal privacy. Several Orwellian references ("*1984*", "*Big Brother [is watching you]*") appeared to be used to accentuate this perception. Many respondents referred to CCTV surveillance as being excessive, intrusive and unnecessary. Its usefulness as a legally enforceable deterrent was also questioned, often with reference to the quality of the evidence it is able to provide: "*I have little confidence in them as to the quality of the output and as to the coverage. They never seem to be on at the right time and/or place*". In terms of trust, concerns about the ownership, access, regulation and control of CCTV data were frequently mentioned, but only by respondents who were otherwise broadly negative about its use.

Car tracking: In response to the question *How do you feel about the possibility of UK road pricing, as charged by the continual tracking of your vehicle?* 84 (60%) of respondents reacted negatively, with 23 (16%) considering such a technology to be positive and 23 (16%) of mixed opinion. The potential for reducing congestion, car dependence and costs (e.g. for light road users) were considered as positive aspects. The use of such a system as a deterrent for criminal acts such as car theft was occasionally offered. However, more than half of the respondents mentioned that such a system was would be unnecessarily expensive and an unwelcome addition to a culture of surveillance that they already considered as being excessive. Comments often made direct reference to invasions of privacy and civil liberty. Trust in the institutions that would manage such a system seemed low. Many respondents appeared skeptical that such a system would be used only for road pricing and many reserved their judgment accordingly: "*If this can be done with suitable anonymity guarantees (which is technically feasible), and it is *only* used for pricing and not policing, I am ok with it.*". Several respondents also indicated a fear that the system could be used to criminalise road users through the assumption that such a system would be able to directly monitor individual vehicle speed.

Respondent consideration of privacy / security issues related to technologies that would utilise knowledge about personal identity

ID Cards: In response to the question *How do you feel about the U.K. Government's proposed ID card scheme?*, around a third (29, 28.9%) of respondents were generally positive. However, over half (84, 58.7%) of the respondents reacted negatively to the proposal. Of the positive comments about ID cards, most were related to their potential as a means of protection against crime (specifically terrorism), though the convenience of centralising multiple forms of ID into one was also occasionally mentioned. However, ID cards were much more frequently considered of as being expensive, unnecessary and ineffective as a means of reducing crime. Of the negative remarks made, ID cards were frequently described in strong terms as a gross invasion of privacy and civil liberty: "*Strongly against. Worried about personal freedoms, data security, general expansion of surveillance*".

Aspects of trust related to the proposed centralising of multiple personal data sources onto one card were oft-remarked upon, as were security risks associated with the potential for data loss / leakage consequent of such activity. Trust in the competence of the U.K. governments ability to protect stored data against abuse and loss was frequently questioned: "*I have very little faith in the government being able to manage all the personal data. Their track record is poor. This data has a potentially high value and will thus be a target for criminals and others who gain value from it*". The cloning of ID cards appeared as being considered inevitable as much as potentially possible, and frequent references were made to several highly publicised media stories in which the U.K. government had apparently lost large quantities of personal data: "*Too much information will be stored and given the Governments recent spate of losing data I would not be entirely trusting of any data protection systems in place*".

Biometric passports: In response to the question *What are your thoughts about the use of Biometric data on U.K. passports*, positive / negative respondent comments were found to be roughly equal (40, 30% positive, 48, 35% negative). Though considered as expensive and difficult to deploy, respondents considered biometrics to offer increased personal security as they were "*hard to fake*" and thus useful for reducing identity theft and the monitoring of migration. However, the vulnerability of data to abuse was also highlighted, with terms such as "*cloning*" and "*hacking*" used to describe such abuse: "*Anything can be copied so it will just be a matter of time until biometric data will be copied too*".

Though issues of privacy invasion and the infringement of civil liberties were mentioned in regard to passport biometrics, the frequency of such comments was markedly less than in the other questions posed: "*Biometric data, in my opinion, is mostly used in terms of identification and not really related to privacy. I am in favour of this as it would help reduce identity theft and other issues regarding identification*". Issues of trust, such as they were commented upon at all, were considered only in terms of security risks related to perceptions of Government ineptitude: "*in light of the number of data cock-ups recently I am not too happy the government will be trusted with this*".

3.4.3 General experience of risk-relevant online service usage

Two specific online activities were considered by the researcher as being of substantial potential risk to the personal privacy and security of their users data. Both activities were chosen based upon the value of such data to criminals and were directly related to financial transactions. These activities were: 1) electronic 'e-commerce' transactions made using credit cards and 2) the accessing of personal banking information through an online service (i.e. *e-banking*).

The vast majority (202, 88%) of respondents reported as having used the WWW to access their personal bank account through an online e-banking service at least once, and almost all (220, 96%) of the respondents reported as having personally made a purchase from a supplier using an online e-commerce website using their own credit / debit card.

Almost all (222, 97%) reported as having made a purchase from a retailer that was previously known to them (e.g. Amazon). Two thirds of respondents (151, 66%) had made a purchase

from a retailer to whom they had no prior knowledge or experience, and 28% (64) had made a purchase directly from an individual who was previously unknown to them (presumably through eBay or a similar 'free ad' type service). The majority (174, 76%) of respondents had however mitigated the risk of their e-commerce transactions failing through use of an online payment gateway service (e.g. 'Paypal') that offered some protection against fraud.

3.4.4 Experience of risk-relevant online service usage in places other than the home

In addition to collecting information about their experiences of e-commerce and e-banking, respondents were also asked *where* and under what circumstances those activities had taken place. For each activity, respondents were invited to select 1) whether they had used their own computing equipment (mobile or otherwise), or equipment to which others had access, 2) whether they had used an access point to the Internet that was private (i.e. their home network), or one that was publicly accessible (e.g. at a Library or Cafe) and 3) whether connection to the Internet was facilitated through a wired or wireless (WiFi) connection. Respondents were able to select as many of the options as they considered appropriate.

General internet browsing

Total counts of the respondents' experience of using the Internet for general web browsing were calculated across the six configurations of wired / wireless connection, public / private network and personal / non-personal browsing device. Percentage distributions of experience for each condition are presented as a graph in figure 3.1.

Given that all 229 respondents completed the questionnaire online, all respondents can be assumed as having experience of using the Internet for general web browsing within at least one of the configurations examined. Within the home environment, the use of wired and wireless Internet connections was found to be roughly equal (192, 83.8% and 191, 83.4% respectively). When using personal browsing devices (e.g. a laptop computer) outside of the home, reports of respondent experience approached the levels seen at home with wireless network connection usage (141, 61.6%), and the use of wireless connections appeared to be much more common than experience with wired connections (86, 37.6%). However, in scenarios where non-personal browsing devices were used, the reverse was found (154, 67.2% wired, 57, 24.9% wireless). This finding was considered as occurring due to respondents considering the browsing devices that were made available to them in their workplace (which were conceivably more likely as using a wired network connection associated Internet access) as falling under the category of 'non-personal' devices.

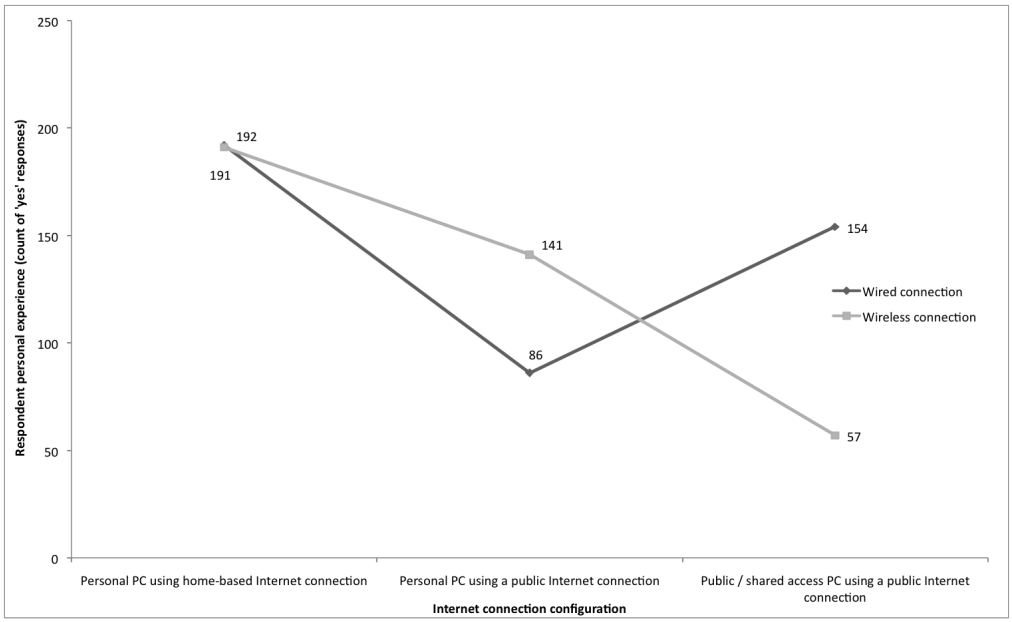


Figure 3.1: Count of respondents ($n = 229$) with experience of general Internet browsing using a personal / non-personal browsing device on a public / private wired or public / private wireless Internet connection.

E-Commerce

Total counts of the respondents' experience of using the Internet to make a purchase from an e-commerce website were calculated across the six conditions of wired / wireless connection, public / private network and personal / non-personal browsing device. Percentage distributions of experience for each condition are presented as a graph in figure 3.2.

Within the home environment, the use of wired and wireless Internet connections was found to be roughly equal (180, 78.6% and 167, 72.9% respectively), and of levels that were comparable to the results found for general Internet browsing. Outside of the home, and when using personal equipment, e-commerce purchases were found to be more common to have been made whilst using a wireless Internet connection (61, 26.6%) than wired (37, 16.2%). However, as seen with general Internet browsing, this finding reversed when using non-personal equipment (39, 17% wired, 10, 4.4% wireless), probably for the same reasons described previously.

E-Banking

Total counts of the respondents' experience of using the Internet to access their personal bank accounts using an e-banking service were calculated across the six conditions of wired / wireless connection, public / private network and personal / non-personal browsing device.

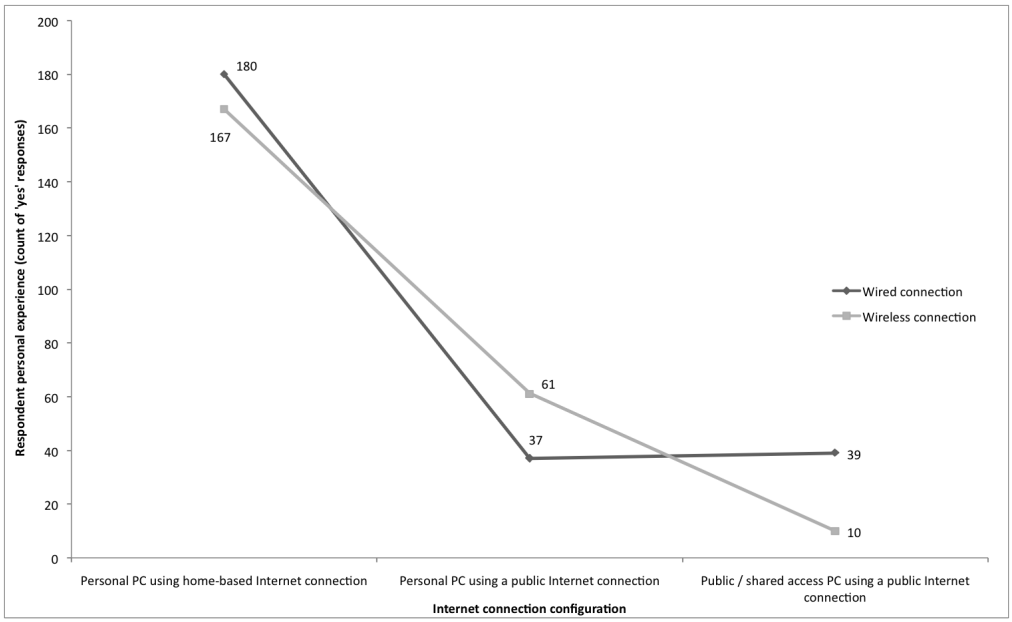


Figure 3.2: Count of respondents ($n = 229$) with experience of making purchases from e-commerce websites using a personal / non-personal browsing device on a public / private wired or public / private wireless Internet connection.

Percentage proportions of experience for each condition are presented as a graph in figure 3.2.

Within the home environment, the use of wired and wireless Internet connections was again found to be roughly equal (165, 72.1% and 153, 66.8% respectively), and of levels that were comparable to the results found for e-commerce and general Internet browsing. Outside of the home, and when using personal equipment, personal bank account access was found to be more commonly made whilst using a wireless Internet connection (52, 22.7%) than wired (32, 14%). Again, as seen with general Internet browsing and e-commerce, this finding reversed when using non-personal equipment (32, 14% wired, 9, 3.9% wireless).

3.4.5 Risk mitigation behaviours in respect of personal technology usage

The results presented in this section were derived from responses to questions that examined the degree to which the respondents made efforts to combat the potential for unauthorized access to their computers as a result of maintaining a connection to the Internet. In terms of how respondents maintained their personal security / privacy on the technologies they used at home, questions were asked in respect of two types of security practice behaviour: **Home network security practises**, including the usage of network communication encryption protocols, particularly with respect to wireless networking (WiFi) and **network communication security practises**, including the monitoring of incoming network communication

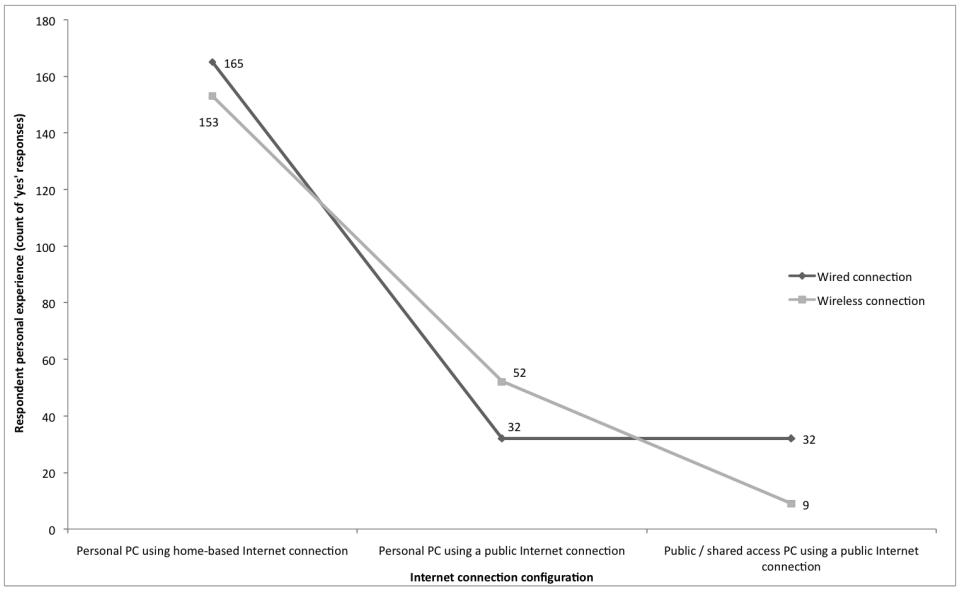


Figure 3.3: Count of respondents ($n = 229$) with experience of accessing personal bank accounts using a personal / non-personal browsing device on a public / private wired or public / private wireless Internet connection.

traffic using security software such as a network 'firewall', and the use of anti-virus / anti-malware software.

Wireless network usage and associated wireless network security practises

In response to questioning about the ways in which they connected to the Internet at home, the vast majority (188, 82%) of respondents reported that they currently maintained a wireless internet connection in their home. 151 (66%) described their wireless networks as being secured from unauthorised usage, either through the use of an end to end communication encryption protocol (e.g. *WEP* [3] / *WPA* [4]), or through limiting access to their networks to specific machines only (e.g. by filtering *MAC* addresses [5]). Only a small minority (25, 11%) of respondents reported as *not* securing their wireless network, but a further 14 (6%) did not know whether their wireless network was secured or not.

[3]WEP: *Wired Equivalent Privacy.*
[4]WPA: *WiFi Protected Access.*
[5]MAC address: *Media Access Control address.*

Usage and understanding of network 'firewalls' as a means of monitoring unauthorised network communications

The degree to which respondents understood the function of a network 'firewall' was captured using an open-ended question: "*What do you understand by the term firewall?*" Almost all respondents stated as understanding the term "firewall" as being referrent to a type of *software* that was involved with *networks* and specifically the *Internet*. When describing the role of firewalls, responses were dominated by terms relating to the provision of two aspects of functionality: 1) security / protection against digitally mounted attack and 2) the provision of network access control.

Security / protection against attack was discussed in terms of the potential that connection to the Internet made the respondents computer vulnerable to the malicious intrusion by unauthorised others. 'Others' were referred to both in terms of people: "*hackers*" and autonomous code: "*Viruses, malware*". "*[A firewall is] a program that maintains the integrity of your system by preventing the invasion or intrusion of malicious or unwanted programs or persons*".

Access control was most frequently referred to in terms of a firewall having the ability to control incoming / outcoming data channels ("*ports*"). The specific abilities of firewalls were discussed in terms of development and enforcement *rules*, *restrictions* and *permissions*: e.g. "*[a firewall] protects computers/networks from unwanted connections based on a set of pre-defined rules*".

Over two thirds of (153, 67%) respondents reported as using firewall software to protect their computers from unauthorised communications made through their Internet connections, of which almost all (142, 62%) could identify the specific brand / type of firewall that they used. However, a third (76, 33%) of respondents did not know whether a firewall was present on their home computer(s) or not.

Personal device operating system and installed software maintenance behaviours

The degree to which respondents made efforts to monitor and maintain the integrity of their personal computer software and operating systems (OS) as a response to sustained exposure to the Internet were examined across a range of items. Such activities included the degree to which explicit efforts were made to monitor and evaluate changes that have been made to their operating system, how Internet 'cookies' / other temporary internet files were managed, and the frequency by which respondents checked their systems for virus / malware infections. Around half (105, 46%) of respondents reported making efforts 'every so often' to maintain their personal computers using dedicated software to assist them in this process. 41 (18%) did so at least once per month, and 30 (13%) reported as doing so at least once per week.

Protection of personal identity online

Data as to how respondents considered the protection of personal identity online as being important to them was first examined by looking at the degree to which respondents chose to use their real names during their online activities (as opposed to a nickname or other pseudonym). The degree to which respondents used their real name whilst using the Internet was measured using scaled response (5-point Likert, [1] never - [5] all the time). Examination of the results yielding an average rating of 3.4, indicating that respondents used their real name on the web only slightly more often than they did not. However, the majority of respondents (192, 84%) also reported as maintaining at least one online pseudonym, with around half (124, 54%) maintaining several.

Further the degree to which respondents engaged with *social networking* websites, where divulgence of personal information is expected (and in many case, at least for practical purposes, necessitated) was also examined. The vast majority (195, 85%) of respondents reported as maintaining a profile on at least one such website (e.g. MySpace, Facebook, LinkedIn). More than half (142, 62%) maintained profiles on two such websites, and roughly one quarter (53, 23%) maintained more than three.

Password behaviour and the protection of personal access credentials on the WWW

To maintain high degrees of user privacy and information security, many online services require their users to generate some form of unique *knowledge-based* identifier such as a username and / or password. The purpose of such an identifier is to allow the service vendor to validate their identity so as to protect their data from unauthorised use. How knowledge-based identifiers, particularly passwords are generated is crucial to their actual security value. Most security guidelines (including the oft-cited *federal information processing guidelines* (1985) [108]) recommend that, to maximise their actual security value, passwords should be both long and constructed of a pseudo-random combination of letters and numbers. Furthermore, passwords are recommended to be changed regularly (and indeed many systems enforce this behaviour), and users are strongly recommended to maintain a different password for each system that they use.

The proliferation of online services in recent years, particularly with regard to e-commerce has led to a substantial increase in the number of passwords that an average Internet user has to maintain. In studies of password usage behaviour, an active Internet user has been found to maintain an average of 25 different passwords at any given time (Florencio & Herley, 2007 [45]), of which some 15 will be required for daily use (Ives et al, 2004 [65]). The requirement to maintain and recall large numbers of passwords however incurs a substantive cognitive burden, and this issue has attracted significant research attention (e.g Adams & Sasse, 1999 [2], Florencio & Herley, 2007 [45]). It is argued for instance that, despite the ideal development of large numbers of unique passwords, the average person will (without training) struggle to maintain more than five passwords at a given time (Adams & Sasse (1999) [2].

Partly as a consequence of this problem, there is a known tendency for users to generate a small number of passwords that are then circulated and / or reused across a range of online services. As variances in the *actual* security of systems that purport to be 'secure' are huge, from a security perspective, such user practice is far from ideal. Further, and as Ives et al (2004) [65] discuss, exposure of a widely used password on a service that is *perceived* to be secure but is in fact not (and is subsequently compromised) can incur a *domino effect* whereby that password is subsequently used to access services with much higher *actual* security measures in place. While suggestions for improvements to password generation that take into account human factors have been offered (e.g. Grawemeyer & Johnson, 2009 [56], Sasse et al, 2001 [128]), despite several decades of research, progress in this area remains slow.

Respondents to the questionnaire were asked whether the number of passwords they maintain in regular use was low or high, and the degree to which they attempted to maximise the entropy or 'strength' of their passwords (e.g. through increased length and / or the combination of numbers and letters). 174 (76%) of respondents reported as actively making efforts to maximise the strength of the passwords that they did generate (e.g. through increasing their length and / or complexity). However, some two thirds (151, 66%) of the respondents also reported as maintaining only a small number of passwords that were repeatedly used across a number of online services.

3.5 Discussion and conclusions

In this chapter, a questionnaire was designed to capture *current* user concerns about privacy and security issues surrounding their current uses of technology. Further, the questionnaire sought to examine the extent to which users of technologies such as the WWW currently manage their personal / private data in response to these concerns. the results of this investigation sought to address **RQ2**: *To what extent do people currently understand the threats and risks associated with the use of ICT services such as the WWW? To what extent does this understanding translate to pervasive / situated computing services?*

From the results of the questionnaire, respondents were found to be far from naive to the potential risks involved in their use of the WWW. Many had personal experience of threats that are commonly associated with use of the Web. Over half (132, 57.6%) of respondents had fallen victim to a virus / malware infection, three quarters 172, 75.1% had personal experience of a 'phishing' type attack, and almost one in five (41, 17.9%) had been the victim of a credit card fraud. On the whole, respondents to the questionnaire presented themselves as being security conscious, and this was reflected in a number of reported behaviours regarding the protection of their personal / private data at home. A substantial number (105, 46%) of respondents made efforts to maintain the integrity of their data and the software that they used on their own computers, even if they did so irregularly. Most (188, 82%) respondents used WiFi to connect to the Internet in their own homes, and two thirds (151, 66%) took steps to secure it against unauthorised usage. Though around a third (76, 33%) of respondents did not use a firewall, the vast majority of respondents understood the types of threat that firewall software is designed to protect them against. However, their security practises regarding password protection was worrying. Despite 174

respondents (76%) reporting as making efforts to strengthen their passwords, 151 (66%) also admitted to maintaining only a small number that were circulated widely.

Personal privacy and identity were considered as being an important factor to respondents, and technologies that involved a perceived invasion of that privacy were carefully weighed up in terms of how they were perceived in terms of their relative cost / benefit. Where a clear benefit could be seen, respondents appeared willing to compromise their privacy in order for that technology to be used. However, respondents were also concerned about who owned, or had access to such data, what could be done with it and how / by whom it was stored. In terms of personal identity, though the use of pseudonyms as a means of protecting personal identity whilst online was common, respondents appeared comfortable with exposing their real names online. Most also appeared comfortable with the use of social networking websites and maintained an account with at least one such service.

3.5.1 Experience of risk-relevant technology / service usage at home and away

The vast majority of respondents to the questionnaire had experience of using technology to conduct activities of financial risk, with some 206 (90%) respondents stating as having used e-banking to conduct their financial affairs online. All respondents (who were old enough to do so) had used the WWW to conduct an e-commerce-based financial transaction using their own credit card. Two-thirds (151, 66%) had done so to make a purchase from a company with whom they had no previous knowledge or experience (with the increased risk that might entail), and potentially riskier still, around one third (64, 28%) had made a purchase from a previously unknown individual.

When using Internet connections outside of the home, network connections tended to be more often wireless than wired when using own equipment, but this appears to reverse when using shared access equipment. Outside of the home, the use of public Internet connections for the purposes of general Internet browsing was found to drop by 26% (191, 83.4% to 141, 61.6%). However, when using public Internet connections for e-commerce and and e-banking, experience levels fell sharply for both: e-commerce (-64%) and e-banking (-66%) (167 [72.9%] to 61 [26.6%], 153 [66.8%] to 52 [22.7%] respectively). Finally, when using non-personal equipment on a public access Internet connection, though 57 (24.9%) stated as having had experience of such a configuration for the purposes of general Internet browsing, only a fraction (10, 4.4%) had used it for an e-commerce transaction and less still (9, 3.9%) for e-banking related activities. Generally speaking, though respondents appeared to be comfortable and experienced in using risk-relevant services using their own equipment and Internet connections at home, these behaviours did not transfer particularly well to their Internet usage outside of the home. On public / shared access Internet connections, though apparently content to use such services for the less risky general browsing, respondents appeared less likely to conduct their e-commerce and e-banking activities. However, it is important to note that the survey did not probe respondents as to why they might be reticent to perform such activities.

3.6 Chapter summary

In this chapter, a questionnaire was developed to investigate the degree to which people currently use technologies that pose a potential risk to the privacy and security of their personal data. In terms of the thesis as a whole, the investigation conducted in this chapter formed the first step towards addressing the extent to which people currently understood the threats and risks associated with the use of ICT services to which they were familiar, such as the WWW (**RQ2**). The questionnaire was designed to probe, covering a broad range of topics that were considered as being relevant to aspects of the less familiar concept of pervasive situated services.

In terms of the degree to which users understanding of privacy / security threats in their current technology use might translate to pervasive / situated computing service usage behaviours, the results of the questionnaire do not yet create a particularly clear picture of the deeper reasoning behind some of the respondents' stated behaviours. However, results did indicate that aspects of trust, privacy and security are important to people, and this appears to be particularly the case with respect to ICT services that are made available outside the more familiar home environment.

From the results of the questionnaire it was found that users are, by and large, highly security conscious and experienced in a wide range of common technology based threats. Further, they were experienced with a number of risk-relevant online activities such as e-commerce and e-banking, but tended to conduct such activities on their home-based Internet connections rather than on public access connections, leaving their use of public access connections for less risky activities such as general Internet browsing.

A question remains as to whether the respondents reticence to engage in riskier ICT usage outside of their home was due to a simple lack of opportunity or to more substantial concerns. In the next chapter, the contributions of the questionnaire in addressing **RQ 2** will be developed further through the design of an empirical research programme that will expose participants in an experimental setting to a common example of a pervasive situated service: a public access WiFi 'hotspot'. The empirical programme will seek to create a type of situation that is pertinent to a known problem associated with pervasive situated services: the need for user-verifiable secure ad-hoc device connection in situations of potential risk. Thus, within the programme, a technology setup that is currently available in the real world (free wifi provision) will be extended to reflect this problem by instantiating several such services in the same physical space. In doing so, participants will need to consider which of several services could be trusted, and which should not be trusted (a situation directly pertinent to addressing **RQs3 & 4**).

Chapter 4

Goals and Methods

4.1 Chapter overview

In chapter one, a number of characteristics of pervasive situated services were presented as being relevant to issues relating to human-computer trust. Those characteristics were: 1) *the wireless / invisible nature of pervasive systems constrain user ability to ascertain the source and intention of the services they provide* and 2) *user trust is often obligated by the requirement to surrender personal information prior to engagement with such systems.* Further, chapter one identified two well-known problems within pervasive computing. These are: *digital service discovery*, i.e. how people come to locate wireless services and *spontaneous secure ad-hoc device association*, i.e. how people come to verify the authenticity of a service and invest their trust in it. Finally, an example of a pervasive situated service that is potentially vulnerable to attacks based upon these two problems has been selected for further investigation: wireless Internet gateway services, or 'WiFi hotspots'.

In chapter two, a questionnaire was developed to probe the degree to which people currently use technologies such as the WWW to conduct activities that pose a potential risk to the security of personal / private information. The results of this questionnaire indicated that people are highly security conscious and careful in the way they conduct their Internet-based activities. Many had experienced instances where their personal data was, or could conceivably been compromised. However, the vast majority of respondents also reported as having had personal experience of using WWW-based services, such as eCommerce and eBanking, to conduct activities that have significant risks attached to them. Importantly however, most only conducted such activities on their home-based networks and equipment. Experience of using public access networks and equipment was found to be low.

In this chapter, the empirical component of the thesis will be presented. The empirical programme will seek to build upon the results of the questionnaire and address the remaining research questions. It will seek to achieve this by creating scenarios whereby participants would be exposed to (and asked to evaluate the trustworthiness of) multiple instances of

wireless pervasive situated services in public space. In doing so, participants will be forced to consider the degree to which they are confident that the services they encounter are truly what they purport to be and, by extension, worthy of their trust.

4.2 Introduction

The first goal of the empirical programme is to draw out and identify those aspects of a person's interaction with a pervasive situated service that are relevant to whether they consider that service to be trustworthy. To address this goal will involve the generation of a deeper understanding about how people currently engage with technologies, and in particular those technologies that require the use of personal and / or sensitive user information. This process will contribute further to **RQ2**, building upon what was learned from the questionnaire reported in the previous chapter. The second goal of the empirical programme is to identify and leverage artefacts in the immediate physical environment that might serve as good evidence that a given wireless service is more likely to be genuine than other similar co-present services that do not utilise that evidence. The overarching goals of the programme are summarised thus:

1. **Explore current user practises**, and identify the context / point at which a persons trust in a service begins to affect their decision as to whether (or not) to accept that service.
2. **Identify evidential cues** in the immediate physical environment that might help or hinder a user to reach a degree of confidence about the trustworthiness of a given service.
3. **Evaluate practical methods by which that evidence can be leveraged** such that the user can make an informed decision about the trustworthiness of a given service.

4.2.1 Design constraints

It is important to reiterate at this point the fact that the research reported in this thesis was conducted within a larger research project (*Cityware*) that involved a number of different stakeholders. A consequence of the author's role as a member of a team working within the *Cityware* project was that certain constraints would be placed upon the way in which the empirical studies reported in this thesis were designed. As such, a number of compromises were made throughout the research process to facilitate the addressing of research goals that were not necessarily directly those of the present thesis.

The most specific instance of this compromise was made manifest in the study reported in chapter five. Within the work package to which the researcher was directly associated, it was considered critical that the work of this thesis would eventually be incorporated into a set of protocols that would specifically seek to solve the problem of secure ad-hoc device

connection within an urban pervasive computing scenario. The design and development of these protocols would (for reasons of timescale) occur concurrently with the research investigations of the author. While some aspects of these protocols would be based upon previous work conducted by Kindberg (as leader of the author's work package), the author would be expected to contribute his findings to as they developed - for instance following the review of literature reported in chapter two.

As a final constraint, it was also considered critical to the wider *Cityware* research group that any protocols that were developed during the project would undergo evaluation by participants within an experimental setting. With limited time available, it was considered preferable to roll together parts of this evaluation into the experiments devised by the author. The knock-on effect of this requirement was that the inclusion of certain experimental materials and data collection strategies would be informed to some degree by other members of the *Cityware* team.

4.2.2 Sampling methods

For the most part, the empirical programme will seek to source a number of its participants through an opportunity-based sampling method. Opportunity based sampling selects participants based upon factors of convenience rather than for their demographic / representative qualities. This sampling method has often been criticised for reducing the degree to which the results that it generates can be generalised to larger populations. The convenience that such sampling offered was considered useful to the researcher for the purposes of increasing participant numbers for subsequent statistical analysis. However, in all of the studies reported, substantial efforts will be made to minimise reliance upon opportunity-based sampling. Wherever possible, a substantial number of the participant pool will be drawn from a cohort of individuals that were specifically employed by *Cityware*. Members of the cohort were selected for their specific demographic qualities (age, gender, occupation etc) to provide for an excellent population sample. For an in depth discussion about the selection and role of the Cityware cohort, see Jay & Stanton Fraser, 2008 [68].

4.3 Roadmap for the programme of research

Revisiting chapter one, the primary prediction of this thesis states: *with no a-priori knowledge of the identity or origin of a given situated service, the decision on whether or not to trust (and thus elect to use) that service will be informed in substantial part by the context and environment in which the situation occurs.* At this juncture it is worthwhile to revisit the research questions that are most directly pertinent to the goals of the empirical research programme:

- **RQ3:** Which aspects of the situational context are considered important to users when they attempt to evaluate the trustworthiness of a pervasive situated service?

- **RQ4:** How do people utilize the aspects identified in **RQ3** to make decisions about whether or not to invest their trust in a given situated service?

Each experiment that the author would devise would seek to present a range of different scenarios to participants. Within each scenario, an instance of a pervasive situated service would be presented. The participant would then be required to evaluate the trustworthiness of that service relative to other co-present services that offered the same potential benefit. The evaluation to be made would be the degree to which the participant felt that the service presented was either genuine (and by extension trustworthy) or potentially bogus / untrustworthy.

Studying trust empirically is not a trivial task to undertake. The development and deployment of research investigations into trust pose a number of challenges for researchers. The creation of methodology that could capture *both* trust-investment behaviour *and* the reasoning behind that behaviour meant that, from the outset, the programme reported here would benefit from collection of both quantitative and qualitative measures to support the author's conclusions. Hence, as far as possible, quantitative measures will be used to compare the effects of particular interventions, while qualitative measures will seek to make targeted inquiries into the nature of the mental models underpinning particular choices made by participants.

4.3.1 Study one

Aims and objectives

The first aim of study one is to create an experimental environment within which a number of co-present wireless digital services could be made available for participants to engage with. As previously mentioned, the form of these services would be wireless Internet gateway services ('WiFi hotspots'). Each WiFi hotspot would, by means of a web-based user interface, require that the participant conduct some form of authentication procedure in order that the service be used. Each individual authentication procedure would be unique, inviting the participant to utilise some aspect of the immediate physical environment to aid them to successfully start using the service.

Unlike typical authentication procedures, whereby the primary role of the process is for the system to authenticate the user, in this scenario, the roles are reversed. Participants will be instructed that at least one of the services that they would come to encounter would have potentially malevolent intentions (and should thus be avoided). During the experiment, each participant will be asked to evaluate each service that they encounter in terms of its trustworthiness, based upon what is involved / presented to them during each authentication process.

At this point, the researcher is concerned more with developing a deeper understanding of the factors that invoke participant discussion related to the potential trustworthiness of the services presented. To this end, the study would not seek to invoke actual trust investment

behaviour from its participants (which would involve some form of simulated risk), but instead concentrate on probing the extent to which a participant states their *intention to trust*, based upon what they encounter. A study involving intention to trust will come later in the programme, based upon further analysis of participant reasoning captured here.

The second aim of the study will be to identify those aspects of the physical environment that might be useful to the participant to engage with the various authentication procedures presented. To achieve this aim will involve a process whereby the research would identify and categorise a number of artefacts in the physical environment, leading ultimately to a conceptual model of how these artefacts might be leveraged to form evidence that a given wireless service is indeed worthy of trust.

Design constraints

In order that the research complement the larger goals of the *Cityware* research project (within which the researcher is based), the design of the experiment would be subject to the following constraints:

1. The experiment must include an example of a novel solution to the problem of *spontaneous secure ad-hoc device association* as an experimental condition. This protocol would be developed by other members of *Cityware*, but involve the author closely - particuarly with respect to the nature of the evidential cue(s) involved.
2. Extending upon the requirement of the first constraint, the experiment must also capture participant reasoning as to the degree to which they perceive each authentication procedure to be *secure*. The term *secure* here refers to the degree to which participants felt confident that an authentication procedure protected them from typical ICT communications threats, such as eavesdropping (e.g. the *man-in-the-middle* type attack).
3. To provide the ability for the *Cityware* team to compare the perceived security value of any novel protocols used, the study would need to include a condition that involved a form of authentication to which the participants would be familiar (e.g. password-based authentication).

4.3.2 Study two

Aims and objectives

Consequent of the findings of study one, study two will seek to develop further a conceptual model of the types of evidence that participants were found to find useful in their reasoning about trust. Dependent upon the degree to which the model develops from analysis of data from study one, this study will seek to either: 1) revisit and potentially rework the evidential cues identified in study one to maximise their effectiveness as a trust-building device, or 2)

investigate other types of evidential cue that were not identified previously by the author, but that emerged through the data collection / analysis conducted in study one.

Design constraints

That the design of this study would be directed by study one, and that participant reasoning about the value of one form of evidence over another might still be unclear demands that this study (again) be limited to measuring *intention to trust*.

4.3.3 Study three

Aims and objectives

By study three, the forms of evidence identified and evaluated in the previous two studies will have coalesced into a conceptual model of evidential cues that can be leveraged to increase the perceived trustworthiness of a given situated service.

Unlike the first two studies therefore, the main objective of study three will be to examine the ultimate effectiveness of one (or more) of these evidential cues in a scenario that involves a real sense of risk to the participants. To achieve this, the design of study three would seek to create a scenario whereby participants would be given the opportunity to choose to invest their trust willingly in a novel situated service that is made available within a public space.

Design constraints

Such that true, voluntary trust investment behaviour be allowed to manifest itself, the design of study three would require that participants believe that their decision to engage with the service presented has certain risks attached to its use. To this end, the design of the study would be such that the researcher would not be present during data capture, and participants would be unaware that they were engaging in an experiment. Such a design has immediate consequences to the amount of reasoning data that the researcher can capture. A discussion of this issue can be found in chapters seven and eight.

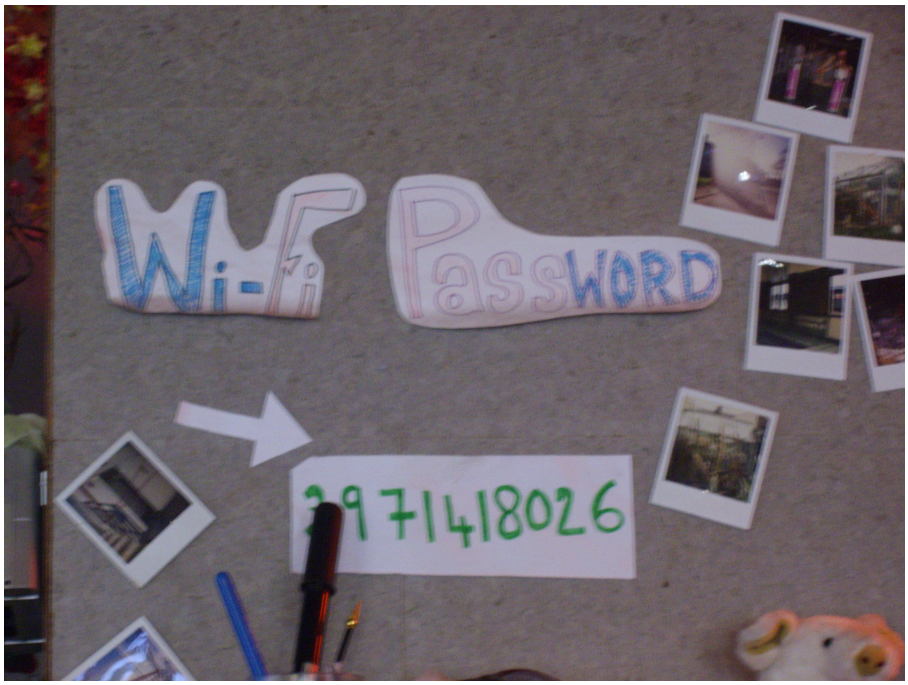


Figure 4.1: Found stuck to the wall of a bar in the city of Bristol, U.K. To which which specific wireless service should this password be considered relevant?

Chapter 5

The effect of physical/digital world linkage evidence as a means of increasing user perceptions of situated service trustworthiness

5.1 Chapter overview

> *Alice is enjoying a coffee in a local cafe. She has her laptop with her and would like to use it to access the Internet if she is able to find a free local wireless Internet service.*
>
> *Scanning for available wireless Internet services, her laptop discovers several networks whose names are all plausible variants on the cafe in which she is sitting. Alice is aware that care should be taken with public unsecured WiFi networks; she might inadvertently connect her computer to a fake network that has been made available by someone who is sitting nearby, or her communications over a bona-fide network may be open to eavesdropping. Either way, information that she considers to be private could be at risk...*

From the scenario described above, and assuming that the management of the cafe do in fact run just one wireless Internet service, the fundamental question asked by this chapter can be considered as thus: *How can Alice be confident that the service to which she chooses to connect is really the one that is offered by the cafe*? This problem, the provision of evidence of a device association is a characteristic of pervasive computing that remains a topic of much research (see section 2.6).

In this chapter, an experiment was designed to investigate how certain *location-based* artefacts in the physical and virtual worlds might be leveraged in order to offer Alice evidence that one wireless situated service - made available amongst a plethora of others - was genuinely provided by the owners of the cafe. In conducting this investigation, the chapter will most directly address **RQ3**: *Which aspects of the situational context are considered important to users when they attempt to evaluate the trustworthiness of a pervasive situated service?* and **RQ4**: *How do people utilize the aspects identified (in RQ3) to make decisions about whether or not to invest their trust in a given situated service?*. However, through providing a means by which the reasoning processes of participants could be captured in a controlled experimental setting, the chapter will also serve to build upon addressing **RQs 1 & 2** as was initiated by the previous chapters.

5.2 How can I be sure this service is genuine? Increasing user perceptions of the trustworthiness of a situated service using location-based evidence

Returning to the introduction to the thesis as presented in chapter one, the principle prediction upon which this thesis is based is that people will seek to utilise evidence that is gleaned from their immediate environment in order to make decisions about whether or not to invest their trust in a given wireless situated service. *Situated* services refer to computing services that are embedded within particular places and contexts. Thus, while the use of such services is achieved without the need for a physical connection such as a cable, the service itself usually bears at least some relevance to the specific place in which it is found. In the most basic sense, the name that is given to the service might be based on the name of the hosting establishment (e.g. '*Starbucks WiFi*').

However, an attacker can create their own situated service with the same (or similarly plausible) name with ease. Thus, as evidence of the degree to which a situated service can be considered genuine, its name alone is of little practical use; it offers little, if indeed nothing at all, in the way of facilitating its discrimination from potentially malevolent services. Moreover, users cannot afford to simply sample each service and see. The act of establishing a full communication connection to a malevolent wireless service may be enough to compromise the security of a user's device and data. This issue, the problem of 'evidently secure device association' (e.g. Kindberg and Zhang, 2003 [76]) is discussed in detail in section 2.6.

5.2.1 Providing links between a digital-world situated service and the physical world to increase user perceptions of trustworthiness

In research conducted into the mechanisms by which e-commerce websites can present themselves as being credible, Fogg (2003) [47] has suggested that a useful cue to the trustworthiness of an e-commerce vendor is the explicit linking of their digital-world website service

to the physical-world company that it represents (see also, Egger, 2001 [41]). The manifestation of such a link can be made very simply through the inclusion of a postal address, land-line phone number or the email address of a company representative within the website. The positive effect that such a simple act invokes is thought to occur because the digital world (the website) is shown as being attached to something in the physical world (e.g. the company headquarters). In providing a sense that the company with whom the user is dealing exists in a tangible physical sense (e.g. in a building), the user is afforded a sense that they are able to mediate the potential risks involved with dealing with their website. Moreover, they are also now equipped with an explicit route of recourse should a transaction with that website fail. With regard to the situated services described in this thesis, if the creation of explicit links between a service deployed on the WWW and the physical world can increase feelings of trustworthiness in risky situations such as e-commerce, might linking a situated service with the immediate physical surroundings of the user provide the same positive effect?

5.3 The design of an experiment to examine the effect of digital-physical world linkage upon user perceptions of situated service trustworthiness

5.3.1 Decomposing the concept of *linkage*: *Physical* and *virtual* linkage

As researchers such as Schilit et al (1994) [130] have observed, a user's *context* at any given time can be understood as being composed to *where [they] are*, *who [they] are with*, and *what resources are nearby*. In the experiment described in this chapter, nearby resources were leveraged to create contextual links between the wireless digital service and the immediate physical world to which that service was made available. Means by which such links could be created have been discussed in the HCI literature, most directly by Barton & Kindberg (2001) [11]. In the experiment presented in this chapter linkages were formed using two methods: *physical* linkage and *virtual* linkage:

***Physical* linkage:** The linking of a digital service to the physical surrounding of the user by using artefacts that are more tightly attached to the immediate physical world (in a simple physical sense) than an intruder could contrive.

***Virtual* linkage:** The linking of a digital service to the physical surroundings of the user by using more interaction between physical artefacts and the digital service than an intruder could contrive.

To evaluate the effect that physical and virtual linkage might have upon user perceptions of the trustworthiness of a given situated service, the experiment described in this chapter was designed based upon a common example of such a service: a public access *WiFi 'hotspot'*.

WiFi 'hotspots'[1] are short range wireless ICT services that typically serve as a means of providing a means of connecting mobile devices to the Internet / WWW. Public access WiFi 'hotspots' are often found in cafes and similar establishments, and are usually offered as a value added service to paying customers. Sometimes a small charge is levied for their use, but frequently the service is offered for free.

5.3.2 Research questions and hypotheses

The experiment sought to evaluate a range of methods that could be used to provide evidence as to the point of origin (and by extension likely 'genuineness') of a wireless situated service (a WiFi 'hotspot'). In doing so, the primary intent of the experiment was to uncover the reasoning processes involved in the ascertaining of situated service 'genuineness' (and by extension trustworthiness), specifically in terms of:

1. How convincing people found the different types of evidence that were provided as they decided whether or not a given service was genuine (addressing **RQ4**).
2. What people thought about the types of evidence that were used with respect to other factors of concern to them, such as convenience and ease of use (addressing **RQ3**).

The principle hypothesis (**H1**) of the experiment was that *user perceptions of the trustworthiness of a given wireless service will be increased by providing a link between that service and the physical surroundings of that user.* Further, with respect to the contribution of physical and virtual forms of linkage, a secondary hypothesis (**H2**) was also formed: *As the strength of the link (measured by the degree to which users perceive that a given link could be plausibly contrived by an attacker) increases, so too will their perceptions of the trustworthiness of that system.*

5.3.3 Experiment environment and materials

To conduct the experiment, a number of independent-yet-copresent WiFi 'hotspots' were created by the researcher (see section 5.4.1) and deployed within a real-world cafe at the campus of the University of Bath, U.K. Though it was acknowledged that the effect of such an environment upon participants responses to the study would be negligible, it was felt that to use a real-world cafe venue would provide some degree of ecological validity, at least above what might be achieved through the use of a more typical laboratory setting. The experiment was conducted while the cafe was closed to the public. To help control for any potential effects relating to prior experience with the existing cafe brand, all visible branding on the walls and tables were replaced with branding of the researcher's own design (figure 5.1), based around a fictional cafe called *Bertorelli's*.

[1] For reference, home-based wireless Internet setups are essentially the same as WiFi 'hotspots'.

Figure 5.1: Examples of *Bertorelli's cafe* branding (leaflet and poster).

5.4 Experimental design

The experiment was a within-subject design that was conducted with the researcher present at all times (i.e. the experiment was an *attended* experiment). Within the experiment, types of evidence as to the genuineness of one service over another were formed from a combination of a web page that was presented to the user (via a laptop supplied) and some physical artefact that was placed within the room.

5.4.1 The independent variable and the creation of experimental materials

Within the experiment, participants were exposed to, and asked to evaluate the *genuineness* of six independent WiFi services that all purported to offer free access to the Internet. The term *genuine* referred to the degree to which participants felt that the service that they were evaluating was bona-fide and provided by the management of *Bertorelli's*. All six services purported to be genuine, and each of the services would use some form of evidence to authenticate itself to the user as being a genuine service offered by the *Bertorelli's* management. The independent variable (IV) was thus *authentication method*, with six levels.

Six conditions of *authentication method* were created by the researcher that systematically manipulated the strength (*fixedness*) of several artefacts that could conceivably offer characteristics of physical and virtual linkage as an evidential cue to their genuineness. To manipulate physical-linkage, three artefacts were chosen as having characteristics of *fixedness* that were considered as being sufficiently differentiated from one another. These were:

Low physical fixedness: Artefacts considered as being of low physical fixedness are mobile artefacts that are accessible by both staff and patrons. Leaflets were considered as offering low fixedness in that they are both loose and of small / less visible physical size.

Medium physical fixedness: Artefacts considered as being of medium physical fixedness are highly visible yet static, accessible by staff but less so by ordinary patrons. A wall mounted poster was selected for this role as it would be clearly visible to both patrons and staff, but less mobile than a leaflet as it would be affixed to the wall.

High physical fixedness: Highly physically-fixed artefacts are highly visible but only fully accessible by staff. A wall-mounted, large form-factor LCD screen was selected for this role as its physical bulk (and wall-securing mechanisms such as drilled-brackets) would make it difficult to access by non-staff.

To manipulate virtual-linkage, where fixedness is measured by the degree to which a user considers the possibility that the interaction it provides could be contrived, three levels were considered. These were:

Low virtual fixedness: Interactions that are considered as being of low virtual fixedness are those that require no *actual* evidence of end-to-end device association. For this role, a simple authentication protocol was developed, whereby a user simply entered a password that is made publicly available somewhere in their immediate physical environment (the acceptance of a password by some unidentified device offering no evidence of the source of that device / communication).

Medium virtual fixedness: Interactions that were considered as being of medium virtual fixedness would involve some evidence of end-to-end device association. Medium virtual fixedness interactions would offer user-verifiable evidence that a connection was being made with the intended service (thus protecting against an *evil-twin* attack), but not necessarily that the connection was secure from eavesdroppers / *man-in-the-middle* based attacks. For this role, *Synchronisation*, a dynamic variant of a *visual hash* (see section 2.6.2) type protocol was developed. In the *Synchronisation* protocol, a simple visual interaction is created between the screen of the user's device and a (highly physically-fixed) public LCD screen in which both screens display a series of protocol-generated images. By comparing the degree to which the images on both screens agree (both in terms of their appearance and the degree to which both screens display the images in sync with one another), users are offered visually-verifiable evidence that the two devices are communicating with one another.

High virtual fixedness: Interactions that were considered as being of high virtual fixedness would involve user-verifiable evidence both of end-to-end device association and communication security (thus offering actual security value against both *evil-twin* and *man-in-the-middle* type attacks). For this role, a multi-stage protocol *Interlock* was developed by the Cityware research group[2]. Within the *Interlock* protocol, a user-verifiable out-of-band communication channel was used to facilitate the comparison of multiple items of user-generated verification information. The multi-part nature of the verification process allowed for a means of detecting devices that were acting as a relay (i.e. a *man-in-the-middle*) through requiring the relay to submit secret information it could not generate, thus exposing itself when the information is finally compared visually by the user. Specific details of the interlock protocol interaction are presented in later in section 5.4.1. Details of the underlying cryptographic technique used in this interaction can also be found in Kindberg, Mitchell, Grimmett, Bevan, and O'Neill (2009)'s: *Authenticating Public Wireless Networks with Physical Evidence* [73].

The relationship between the levels of physical and virtual linkage for the six conditions generated for *authentication type* are presented as a matrix in table 5.1 below:

[2]While the author was not involved in the specific cryptographic mechanisms that underpin the Interlock protocol, he was directly involved in the development of the user experience aspect of the protocol's design.

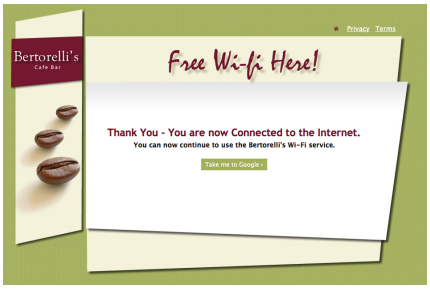
	Physical-linkage *fixedness*		
	Low	*Medium*	*High*
Virtual-linkage *fixedness*			
Low	Password on leaflet	Password on poster	Password on screen
Medium	N/A	N/A	Synchronisation
High	N/A	N/A	Interlock

Table 5.1: Conditions of the independent variable *authentication type*. N.B. The sixth condition, a control condition that did not utilise physical or virtual linkage, is not represented in the table.

A total of six independent WiFi 'hotspots' were created, each of which was named (via the SSID that is broadcast by each service) using a plausible variant of the *Bertorelli's* brand name. For each service, an associated authentication methods was designed, and the specifics of each authentication process are presented in turn:

Direct Connection (SSID: 'Bertorelli1')

The *direct connection* condition (figure 5.2) had no authentication protocol assigned to it and was used in the experiment as a control condition.



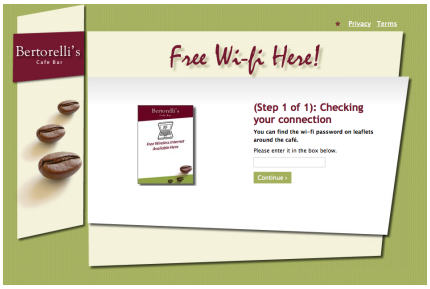
(a) *Direct connection*: Authentication screen (displayed on participant's laptop screen)

(b) *Direct connection*: Connection complete (displayed on participant's laptop screen)

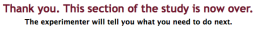
Figure 5.2: *Direct connection*: Authentication procedure (connect to Internet service)

Password on leaflet (SSID: 'Bertorelli2')

In the *password on leaflet* condition (figure 5.3), participants were required to enter a password that could be found on any of a number of leaflets littered on tables throughout the cafe space. Upon locating and entering the password, the participant was instructed to click a button labelled 'continue'.



(a) *Password on leaflet*: Authentication screen (displayed on participant's laptop screen)

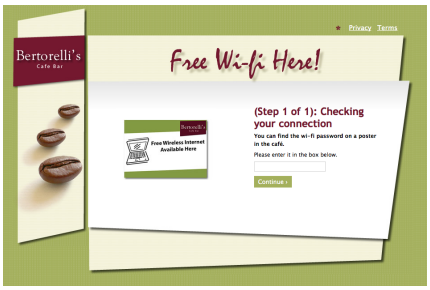


(b) *Password on leaflet*: Connection complete (displayed on participant's laptop screen)

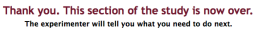
Figure 5.3: *Password on leaflet*: Authentication procedure (connect to Internet service)

Password on poster **(SSID: 'Bertorelli3')**

In the *password on poster* condition (figure 5.4), participants were required to enter a password that could be found on a large poster affixed to the cafe wall. Upon locating and entering the password, the participant was instructed to click a button labelled 'continue'.



(a) *Password on poster*: Authentication screen (displayed on participant's laptop screen)

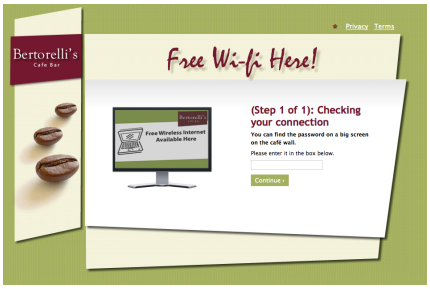


(b) *Password on poster*: Connection complete (displayed on participant's laptop screen)

Figure 5.4: *Password on Poster*: Authentication procedure (connect to Internet service)

Password on screen **(SSID: 'Bertorelli4')**

In the *password on screen* condition (figure 5.5), participants were required to enter a password that could be found on a large LCD screen that was affixed to the cafe wall. Upon locating and entering the password, the participant was instructed to click a button labelled 'continue'.



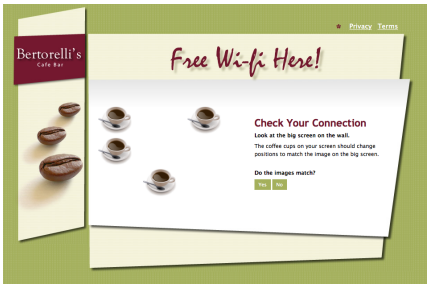
(a) *Password on screen*: Authentication screen (displayed on participant's laptop screen)

(b) *Password on screen*: Connection complete (displayed on participant's laptop screen)

Figure 5.5: *Password on Screen*: Authentication procedure (connect to Internet service)

Synchronisation **(SSID: 'Bertorelli5')**

In the *synchronisation* condition (figure 5.6), participants were required to observe and judge the synchronicity of a dynamic sequence that appeared on both the large LCD screen (as used in the *password on screen* condition) and their laptop screen. An apparently random array of 'coffee cups' was used to form the sequence, and the configuration of the array changed at a rate of once per second. Upon satisfaction that the two displays were synchronised, the participant was instructed to click a button labelled 'continue'.



(a) *Synchronisation*: Authentication screen (displayed on participant's laptop screen)

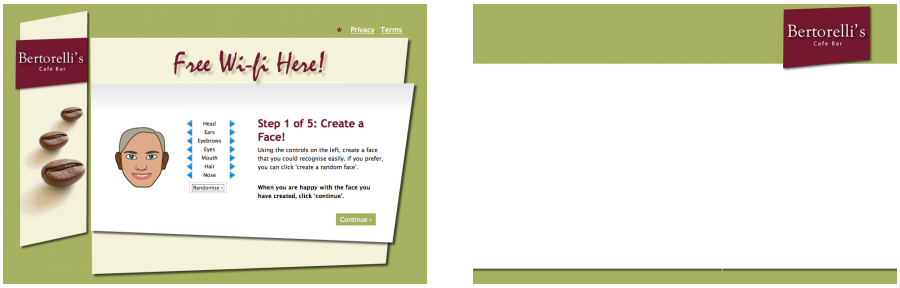
(b) *Synchronisation*: Synchronised display (displayed on public LCD screen)

Figure 5.6: *Synchronisation*: Authentication procedure (connect to Internet service)

Interlock **(SSID: 'Bertorelli6')**

In the *Interlock* condition, participants were required to engage in a multi-part authentication procedure conducted between the participant's laptop and the large LCD screen. At the core of the Interlock protocol is the selection (by the participant) of two items of information that are sent wirelessly to the system that controls the LCD screen. A successful authentication is achieved when the participant judges both items of information, as subsequently displayed on the LCD screen, are the same as the items they originally chose.

Stage 1 - Choose a face: The participant is invited to create their own unique avatar (a face) through the manipulation of a number of parameters including the eyes, mouth and nose (figure 5.7). At this stage, the large LCD screen remains largely blank. Upon selection of a unique avatar, the participant is instructed to click a button labelled 'continue'.



(a) Choose a face: Laptop screen) (b) Choose a face: Public LCD screen)

Figure 5.7: *Interlock* procedure stage 1: Choose a face

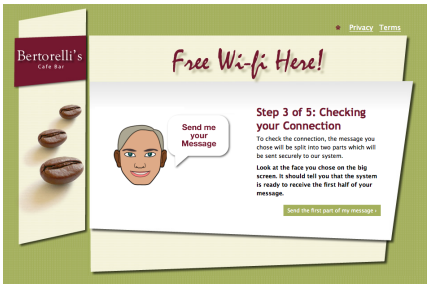
Stage 2 - Choose a phrase: Having chosen their 'face', the participant is then invited to choose a short phrase, as made available by a drop-down list (figure 5.8). At this stage, the large LCD screen remains blank. Upon selection of a phrase, the participant is instructed to click a button labelled 'continue'.

(a) Choose a phrase: Laptop screen) (b) Choose a phrase: Public LCD screen)

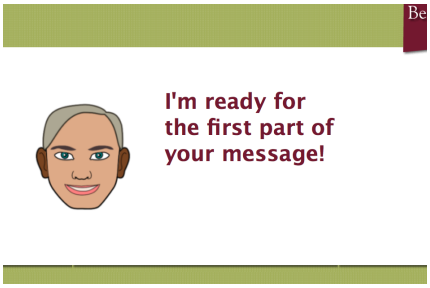
Figure 5.8: *Interlock* procedure stage 2: Choose a phrase

Stage 3 - Send the first part of the phrase to the system: Having now chosen both 'face' and 'phrase', the participant is told that the system will now attempt to *check their connection*. In order that this be achieved, the following explanation is given: "*to check the connection, the message you chose will be split into two parts which will be sent securely to our system*". The participant is then instructed to look at the large LCD screen, whereupon the 'face' that they chose would ask them to "*send the first half of the message*" by clicking the similarly labelled button on their laptop screen (figure 5.9).

Stage 4 - Send the second part of the phrase to the system: Upon the sending of the first half of the message, the LCD screen acknowledges receipt of the message to the user, and the participant is then instructed to send the final part of the message.



(a) Check the connection, step 1: Laptop screen)



(b) Check the connection, step 1: Public LCD screen)

Figure 5.9: *Interlock* procedure stage 3: Check the connection

Figure 5.10: *Interlock* procedure stage 3.1: Message received

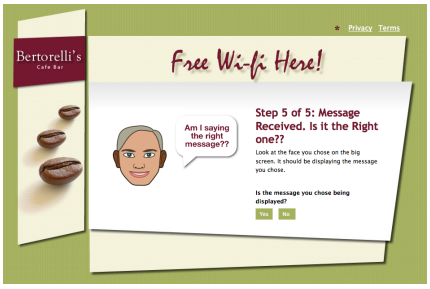
(a) Check the connection, step 2: Laptop screen)

(b) Check the connection, step 2: Public LCD screen)

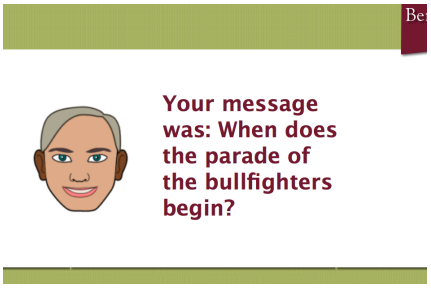
Figure 5.11: *Interlock* procedure stage 4: Check the connection

Stage 5 - Is it the right message? Upon receipt of the final part of the phrase, the avatar on the large LCD screen displays the complete phrase. At this point, the display on

the laptop asks the participant whether the phrase shown on the LCD screen is the same as the phrase that they originally chose. The participant is instructed to select 'yes' or 'no' as appropriate (figure 5.12), upon which the authentication process is concluded.



(a) Do the phrases match?: Laptop screen)



(b) Do the phrases match?: Public LCD screen)

Figure 5.12: *Interlock* procedure stage 5: Compare phrase sent from laptop with phrase displayed on public LCD screen

5.4.2 Dependent variables / experimental measures

The *Bertorelli's* experiment consisted of two phases. Each participant completed both phases consecutively. In the first phase (evaluation phase), each participant was exposed to each of the six WiFi hotspot services one-by-one and, through completion of the authentication mechanism provided by each service, were asked to evaluate them in terms of their 'genuineness'. In the second phase (post-evaluation interview), participants were invited to discuss their experiences with all of the services that they had encountered in a semi-structured interview.

As a precautionary measure inspired by Kindberg et al, 2004 [75], explicit mention of the terms *trustworthiness* and *trust* were not made either at the outset of the experiment, nor during the evaluation phase of the study. The rationale of this was to avoid potentially leading participants into specific concerns about the role of trust, and instead to allow the concerns of the users to be elicited "naturally". Specific questions about the degree to which the participants considered the services as being *trustworthy* were posed to them only during the post-evaluation interview.

During the evaluation of each service encountered, how the participants considered each system that they encountered as being *trustworthy* was instead indirectly measured during the experiment through measure of *confidence*: i.e. *'how confident were the participants in each of the access points they had encountered'?* The measure of confidence in any of the systems had two components (not independent): 1) whether the participants thought that each particular service was being provided by *Bertorelli's* and 2) Whether they thought that that service was vulnerable to attack by a third party. The primary experimental measure was therefore *confidence*, and this was broken down to form three dependent variables (DV's). These were: 1) *confidence in the genuineness of the service*, 2) *confidence in the trustworthiness of the service* and 3) *confidence in the security of the service.*

In both phases of the experiment, measures for each DV were collected through a combination of quantitative (direct question) and qualitative (free-form verbal reasoning) response. Measures for DV1 were obtained in both evaluation phase and the post-evaluation interview, whereas direct measures for DVs 2 & 3 were obtained only during the post-evaluation interview.

5.4.3 Experimental procedure

Prior to each trial, the experimental space was set up with several interaction props placed within the cafe. These were:

1 large form-factor LCD Screen (42") The screen was fixed to the wall of the cafe using brackets. The screen displayed content only during times when a particular experimental condition required it to do so. At all other times, the screen was blank.

One large paper poster (A1 Size, landscape) The poster was printed at the same dimensions as the LCD screen and was affixed to the wall adjacent to the screen.

8-10 Leaflets (A5 size) Leaflets were littered across tabletops throughout the space.

A WiFi enabled laptop computer A Toshiba laptop (operating system: Microsoft Windows XP, Internet browser: Mozilla Firefox) was set up on a table in the cafe for the participants use during the evaluation phase and post-evaluation interview.

Hardware to support six independent wireless networks / WiFi 'hotspots' The six wireless access points and their associated web-servers were concealed behind the cafe bar away from view. Access to the LCD screen was supplied through VGA cable running behind the bar.

As previously discussed, each experimental trial was conducted in two phases. Phase one was the main evaluation phase (duration 20-30 mins) during which participants were exposed to each of the six WiFi 'hotspot' services and their corresponding authentication procedure. Phase two was a post-evaluation semi-structured interview (duration 20-30mins). The duration of each complete experimental trial was thus between 40 minutes and one hour.

Experimental phase

Participants performed the tasks set out by the experiment as individuals. Each participant was first met by the researcher outside of the cafe space, and their written consent and demographic details were obtained (see appendix B, section B.1). The participant was then invited to read an instruction sheet that explained the nature of the experiment and activities that it would entail (see appendix B, section B.2). The instructions also provided a short scenario that served to inform the participant as to the potential threats involved in un-secured / public access WiFi use, and to explain the degree of help they could expect

to receive from the researcher during the study. The scenario presented in the instructions was as follows:

> Imagine you are visiting a cafe to try to get an Internet connection. You have your laptop computer with you, and you know that *Bertorelli's* cafe happens to have a wireless Internet access point that you can use during your visit.
>
> When you sit down at a table and try to connect to the wireless network you see that there are six different wireless networks that appear to be provided by *Bertorelli's* cafe. You know that at least one of these is genuinely provided by *Bertorelli's* cafe, but one or more may be from another source pretending to be *Bertorelli's* cafe. There are two ways in which wireless connections are susceptible to attack:
>
> 1. Somebody might 'listen-in' to the wireless communications made between your laptop and a genuine access point to the Internet such as the one provided by *Bertorelli's* cafe.
> 2. Somebody might have created an entirely fake network to which you can still connect, potentially giving away passwords or other information and inadvertently giving the faker access to your computer.

That the participants were primed as to the specific threats involved in public access WiFi usage was considered important to their eventual evaluation of the protocols they would encounter in terms of their security value. Participants were informed that the primary task of the study was to evaluate each of six WiFi 'hotspot' services that they would encounter during the study, based upon the knowledge that only one service was *definitely* genuinely provided and sanctioned by the *Bertorelli's* cafe venue (i.e. the other five may, or may not be genuine). The participant was then guided into the cafe space and was invited to take a seat at a table with a laptop computer. The laptop was powered up and ready for immediate use though not connected to any service. The researcher demonstrated the procedure of searching for, connecting to and disconnecting from a wireless networking service using the Windows XP operating system (an existing University of Bath wireless network facility was used for this exercise). Only after the participant was happy with managing this procedure on his or her own was the experiment allowed to start.

The six methods of connecting to WiFi services that were supposedly provided by *Bertorelli's* were then completed at the participant's own pace. Connection was achieved in each instance using a corresponding web-based user interface displayed on the user's laptop. Each connection attempt was initiated with the participant being presented with the introductory *splash screen* presented in figure 5.13.

The order that the six services were encountered was supplied in pseudo-random order by the researcher. Each trial was completed when the participant had successfully negotiated the authentication process presented by the service, at which point a webpage informed the participant that the trial was complete. Throughout the evaluation phase, two researchers were seated with the participant. One researcher encouraged the participant to discuss their thoughts regarding the genuineness of each of the wireless services that they encountered (i.e. to "think aloud"). The other researcher took notes. An audio-visual record of the entire session was taken using camcorder.

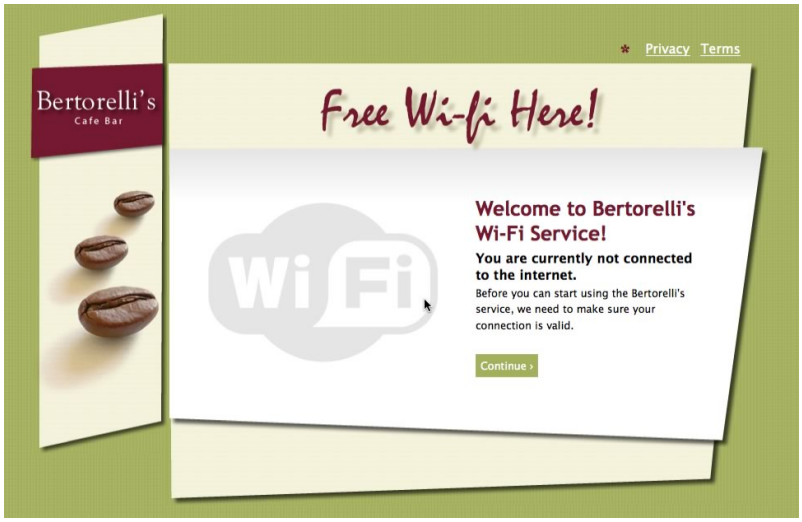


Figure 5.13: *Bertorelli's cafe* WiFi service 'splash' screen.

After the authentication mechanism of each service had been completed, participants were asked to assign a score to that service, based on responses to the following two questions (both 6-point Likert scale):

1. Was this wireless network provided by *Bertorelli's* Cafe? ([1] fake – [6] genuinely provided by *Bertorelli's*).
2. How easy was it to connect to this wireless network ([1] very easy – [6] very difficult).

Post-experimental interview

It was expected that the reasoning underlying each participant's decision as to their confidence with one service over another would be multi-faceted, including aspects of each participant's existing mental model of computer / WiFi security and the impact (if any) that the inclusion of physically-linked artefacts / virtual-linked interactions had upon that model. As it was considered likely that other factors that could not be identified and / or anticipated in advance could form some of this reasoning, the post-evaluation interview strategy allowed for such reasoning to be elicited both spontaneously by the participant and by direct probe by the experimenter.

Thus, immediately after completion of the evaluation phase, a semi-structured interview was conducted with each participant, based around a set of 10 questions that were presented in random order. Five of the questions involved rank ordering the six WiFi 'hotspot' services that they had encountered against some range of criteria (e.g. "fake" to "genuine") and the other five involved assigning each of the six services to a simple binary choice of yes / no (e.g. "I would use this network to check my email"). A summary of the questions posed / response type used is presented in table 5.2:

#	Question	Values
1	Using the terms *insecure* to *secure*, how would you rank the access points you have tried today?	Insecure - Secure
2	Using the terms *easy* and *difficult*, how would you rank the access points you have tried today?	Difficult to complete - Easy to complete
3	Using the terms *confusing* and *straightforward*, how would you rank the access points you have tried today?	Confusing – Straightforward
4	Using the terms *genuine* and *fake*, how would you rank the access points you have tried today?	Fake (not provided by Bertorelli's - Genuine (provided by Bertorelli's)
5	Using the terms *untrustworthy* and *trustworthy*, how would you rank the access points you have tried today?	Not Trustworthy – Trustworthy
6	I would use this wireless connection to *read an online newspaper*	Yes / No
7	I would use this wireless connection to *check a train time*	Yes / No
8	I would use this wireless connection to *check my online email account*	Yes / No
9	I would use this wireless connection to *buy something from an online shop using my credit card*	Yes / No
10	I would use this wireless connection to *access my bank account online*	Yes / No

Table 5.2: Post-evaluation interview questions. Each question was prefixed with the instruction: "For each system, please drag the appropriate card to answer the question...".

Figure 5.14: Interview question interface.

Collection of quantitative measures for each question were collected using a graphical web-based system that was developed by the researcher (figure 5.14). The system allowed participants to make their response selections by clicking and dragging icons that represented each of the six services around the screen.

The five binary choice questions related to the degree to which our participants understood and considered the authentication methods presented to them by each system as being sufficient in terms of risk investment over a number of different real-world transaction scenarios. Five general transaction types were explored, ranked in order of their actual risk to personal information (low to high) in table 5.3:

Internet activity / behaviour	**Required personal data involvement**	**Actual threat potential**
Read an online newspaper	None	Limited, though connection could facilitate malware.
Check a train time	None	Limited, but could indicate user travel plans.
Access a personal email account	Username / password	Potentially significant, dependent upon value of email account.
Make a transaction using own credit card	Username / password / credit card details	Significant, though compensation for online fraud is well established and losses are limited by credit limit of card.
Access personal bank account online	Username / password / additional security information	Significant. Potential for significant financial loss if account is compromised.

Table 5.3: Summary of online risk behaviour scenarios examined. Risk levels associated with each scenario increases low-high.

During the interview, participants were encouraged to 'think aloud' as they answered each of the 10 questions posed and the researcher probed the participant for more information on points raised during this process. At the end of the interview, the researcher asked the participant to explicitly choose one of the six networks as being the one they considered to be the genuine *Bertorelli's* WiFi 'hotspot' service.

5.4.4 Participants

The participants were 28 individuals (M=20 [71%], F=8 [29%], modal age range = 26-30) recruited by email invitation and general opportunity sampling. 17 (61%) of the participant pool were sourced from staff and students within the University of Bath, with the remaining 11 (39%) sourced from the *Cityware* project cohort. The *Cityware cohort* was a diverse group of non-student residents of the City of Bath, U.K. who were employed as part of the University of Bath led *Cityware* research project to assist in studies carried out by its researchers (e.g. Jay & Stanton Fraser, 2008 [68]). Participants were rewarded for their time with a small treat (a hot drink and some biscuits). Ethical approval for the study was applied for and granted by the University of Bath department of Psychology and consent to participate was obtained from each participant in writing (see appendix B, section B.1).

5.5 Results

5.5.1 Trust measures in response to specific use-case scenarios

Categorical (yes / no) response data was collected for each of the six services evaluated based on whether participants would consider trusting using those services to perform a set of five specific online behaviours (see section 5.4.3). Mean acceptance rates (% count of positive 'yes' ratings) were calculated for each service / behaviour examined and are presented as a table in figure 5.4.

On first examination, the risk acceptance rates across the five scenarios examined tended to confirm that the author's rankings of the risks associated with those scenarios mirrored those of the participants. As the risk associated with the scenario increased, there was a corresponding decrease in acceptance rates and this was found to be largely independent of the service / authentication methods evaluated.

Reading a newspaper and checking a train time

"*With an online newspaper there's no risk ...It's less risk, I wouldn't be that scared*". Low risk activities such as the reading of an online newspaper and the checking of a train time were considered as being a largely acceptable activity to perform on a public WiFi service, independent of the specific service that was used. In the newspaper / train time scenarios,

	Direct	Leaflet	Poster	Screen	Syncn	I/lock
Read an online newspaper	19	22	26	24	26	24
	(67.90%)	(78.60%)	(92.90%)	(85.70%)	(92.90%)	(85.70%)
Check a train time	20	23	26	27	27	25
	(71.40%)	(82.10%)	(92.90%)	(96.40%)	(96.40%)	(89.30%)
Access an email account	7	11	13	15	20	20
	(25.00%)	(39.30%)	(46.40%)	(53.60%)	(60.70%)	(71.40%)
Make a credit card transaction	0	0	0	0	9	9
	(0.00%)	(0.00%)	(0.00%)	(0.00%)	(32.10%)	(32.10%)
Access an online bank account	0	0	1	1	6	5
	(0.00%)	(0.00%)	(3.60%)	(3.60%)	(21.40%)	(17.90%)

Table 5.4: Risk acceptance rates for five online activities, listed in ascending degrees of risk. The conditions are listed in ascending degrees of fixedness (left to right), starting with physical linkage and followed by virtual linkage.

only the idea of performing such activities with the direct connection protocol reduced the acceptance rate to below 70%. Participants generally perceived little to no risk to be associated with either activity as no personal information was understood by them as being required to use such a service: "*I haven't given out any personal information. If someone goes to a lot of trouble finding out who I am to find out which paper I read, they can have that [information]*".

Accessing a personal email account

The accessing of a personal email account was considered acceptable by around half (46.4%) of the participants if using a service that employed a minimum of medium physical fixedness as an authentication measure. Acceptance rates were however found to increase with the virtual-fixed services, peaking at 71.4% for the highly-virtually fixed *Interlock* protocol.

That the accessing of an email account would require the use of personal information, specifically a username and / or password, was frequently mentioned. Participants were reticent to expose such information without some assurance as to the security of the network service they were using: "*Anything that I have which has a password on, I wouldn't do it on a network where I wasn't confident of the configuration.*" With regards to the physically-fixed protocols examined, some participants would trust them enough to use them to access their email account, but only in times of real need: "*Depends on how urgent it was. If it was urgent I'd probably use any of them - if it wasn't urgent, I probably wouldn't*". With the virtually-fixed protocols however, this consideration was relaxed for those participants who considered them as offering assurances as to their security: "*I think that they are secure, they are trustworthy because the level of difficulty for being hacked is high... high enough to be secure and trustworthy.*"

Credit card usage and the accessing of personal bank accounts

The use of personal credit cards and the accessing of personal bank accounts was universally rejected by the participants in all authentication methods that did not utilise some form of virtual-fixedness based evidence in their design. However, even with the virtually-fixed protocols, there appeared to be little difference between acceptance rates for the medium and high virtual-fixedness, and no scenario / service configuration ever exceeded 33% acceptance.

Regarding the use of credit cards, participants were generally opposed to the idea of using any public network connection: "*I wouldn't. I'd only use my connection at home*", though some would do so if there was some necessity to do so: "*It depends how badly I need to buy something right now. I would avoid using it unless I had to*". As with the accessing of email accounts, the lack of perceived control and assurance with regards to the security of the network was oft-mentioned: "*I wouldn't be sitting there keying in my credit card. I would feel all the security aspects are locked down to me*". One participant also mentioned discomfort about the nature of wireless connectivity: "*Because the wireless [signal] goes everywhere, it feels less secure*".

E-banking was generally considered as being more risky than credit card usage, and the lack of explicit risk-mediating devices that e-banking services offer relative to Internet credit card use (such as payment protection) were occassionally mentioned as an explanation for this difference: "*I guess this would be [more risk] than credit cards. I probably wouldn't do either if there wasn't a secure connection. At least with credit cards you've got some kind of protection*". However, the rejection of using a network service outside of the home to conduct e-banking activities was, like credit card usage, generally couched in a perceived lack of control over communication security: "*I would never use a wireless network for on-line banking. Or a wired network in an internet cafe, because I haven't configured the network myself*".

5.5.2 Participant confidence in service *authenticity*

The main quantitative measure of DV1, *confidence in the authenticity of the service* was participant response to the question "was this wireless network provided by *Bertorelli's* Cafe?", which was obtained through a six point Likert scale ([1] fake – [6] genuinely provided by Bertorelli's). This question was posed on two separate occasions: immediately after exposure to each individual service / condition (single-condition score), and additionally at the end of the main evaluation phase, where each participant was asked to rank order all of the six conditions (cross-condition score). Examination of both sets of results allowed for a measure of participant consistency: i.e. whether they had changed their opinion of a particular condition later, having been exposed to all six conditions. Results for each set of responses are examined in turn.

Confidence scores for service genuineness: *Single-condition* and *cross-condition* scores

Whereas the single-condition scores were on a six-point Likert scale, the cross-condition scores were generated using the interactive tool described in section 5.4.3. Using the tool, participants manipulated icons that represented the six networks to obtain their desired ranking along a horizontal axis, resulting in a range of values from 0 to 1300 (see figure 5.14). Prior to analysis, any two icons that overlapped with one another were adjusted to be equal at their mid-point.

The single-condition and cross-condition scores were independently analysed across all six conditions using a one-way repeated measures analysis of variance (ANOVA). For the single-condition scores, a significant main effect was observed of *authentication method* [Wilks' Lambda = .25, F(5,23) = 13.619, $p<0.01$], and a multivariate partial eta squared value of .75 suggested a large effect size. Similarly, for the Cross-condition scores, ANOVA showed a significant main effect for *authentication method* [Wilks' Lambda = .27, F(5,23) = 12.70, p ¡ 0.01]. A multivariate partial eta squared value of .79 suggested a large effect size.

Pairwise comparisons (repeated measures t-tests) were then performed on both single- and cross-condition scores between successive condition pairs (in which first physical linkage and then virtual linkage increased). The resulting significant pairings and mean scores are presented as a graph in figure 5.15. To enable better comparison, both sets of scores have been normalised.

5.5.3 Participant confidence in service *trustworthiness*

The main quantitative measure of DV2 *confidence in the trustworthiness of the service* were responses to the question "Using the terms *untrustworthy* and *trustworthy*, how would you rank the access points you have tried today?". Confidence scores for service trustworthiness were obtained only in the post-evaluation phase of the experiment, thus the figures reported here are the cross-condition scores.

Repeated measures ANOVA indicated a significant effect for *authentication method* across scores for *trustworthiness* [Wilks' Lambda = .21, F(5,23) = 17.77, p<0.01]. Multivariate partial eta squared values of .79 suggests a large effect size. Pairwise comparisons (repeated measures t-tests) were performed on the scores between successive condition pairs in which first physical linkage and then virtual linkage increased. The resulting significant pairings and normalised mean scores are presented as a graph in figure 5.16.

5.5.4 Participant confidence in service *security*

The main quantitative measure of DV3 *confidence in the security of the service* were responses to the question "Using the terms *insecure* to *secure*, how would you rank the access

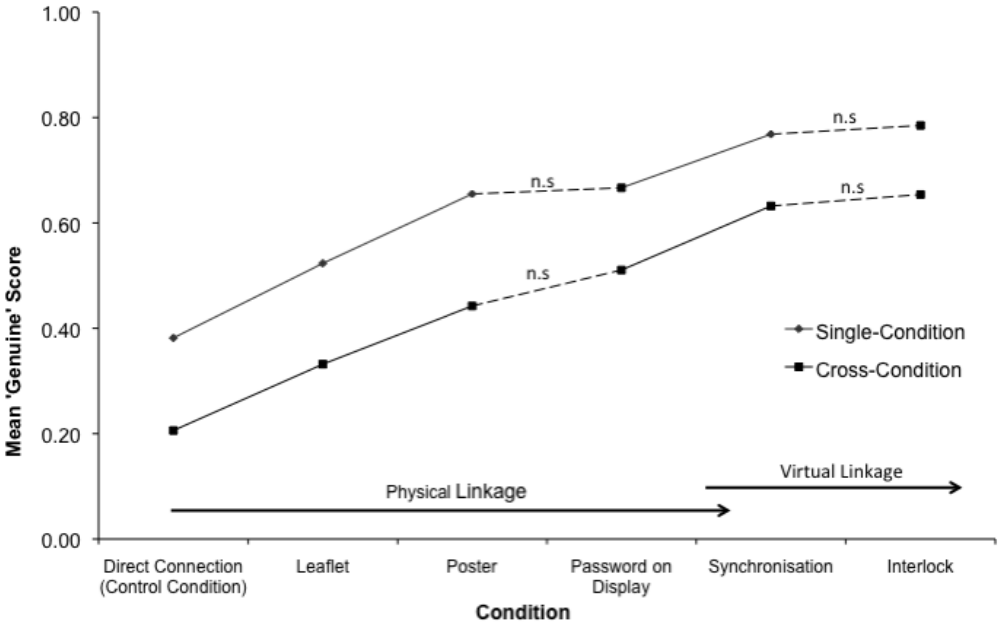


Figure 5.15: Confidence in service 'genuineness' across the six conditions of authentication method. Raw scores converted to log values for comparison. None significant pairwise comparisons (t-test) are highlighted with a dashed line.

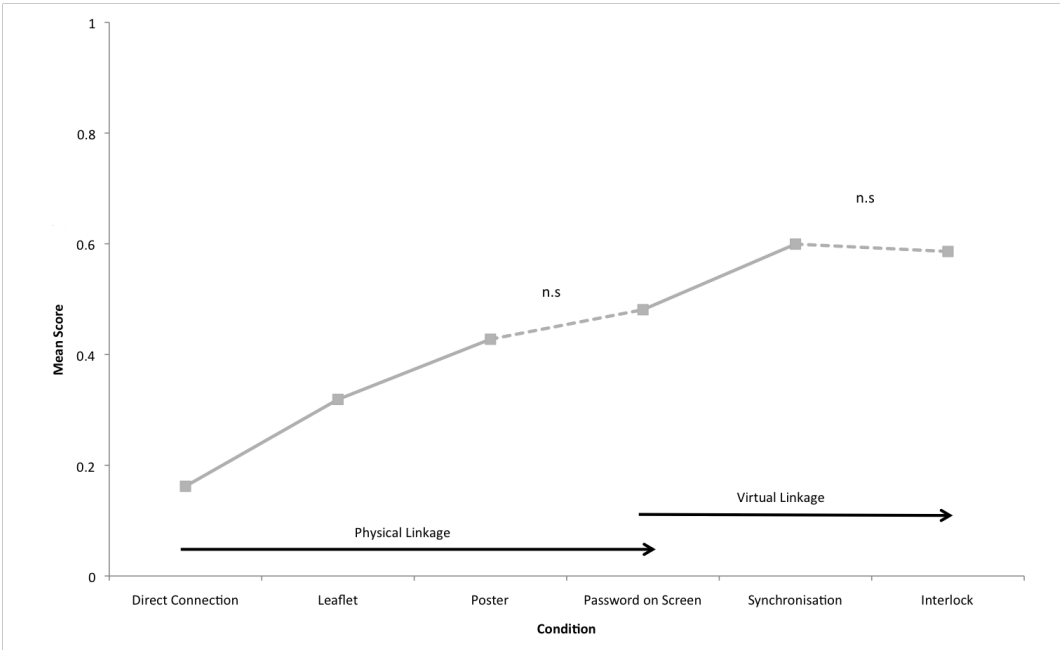


Figure 5.16: Confidence in service 'trustworthiness' across the six conditions of authentication method. None significant pairwise comparisons (t-test) are highlighted with a dashed line.

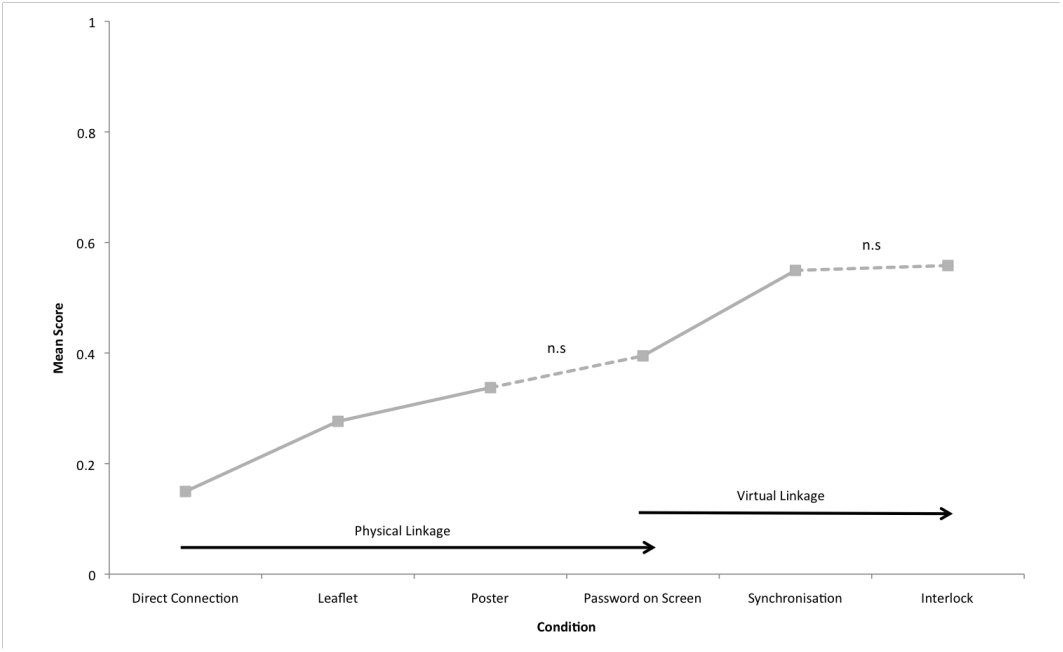


Figure 5.17: Confidence in service as being 'secure' across the six conditions of authentication method. None significant pairwise comparisons (t-test) are highlighted.

points you have tried today?". As with DV2, confidence scores for service security were obtained only in the post-evaluation phase of the experiment, thus the figures reported here are cross-condition scores.

Repeated measures ANOVA indicated a significant effect for *authentication method* across scores for *secure* [Wilks' Lambda = .24, $F(5,23) = 14.61$, $p<0.01$]. A multivariate partial eta squared value of .76 suggested a large effect size. Pairwise comparisons (repeated measures *t*-tests) were performed on the scores between successive condition pairs in which first physical linkage and then virtual linkage increased. The resulting significant pairings and normalised mean scores are presented as a graph in figure 5.17.

5.5.5 Participant reasoning about the impact of physical and virtual linkage upon perceptions of service genuineness, trustworthiness and security

During both evaluation and post-evaluation interview phases, participants were encouraged to 'think-aloud' and to vocalise their thoughts and reasoning freely as they completed the experimental exercise. The audio records of the 'think-aloud' protocols for 27 participants[3] were transcribed and analysed using a form of protocol analysis. The specific protocol analysis technique used in the experiment was a variation of the *Grounded Theory* methodology.

[3](28 participants were involved in the experiment, but 1 video recording was mislaid).

Grounded Theory is a systematic methodology for the analysis of complex written content without a prior set of assumptions or hypotheses. Using this approach, common themes emerge through a process of open coding; that is to say that the coding scheme was developed in response to the data itself, as opposed to being generated prior to examination of the data. As it was expected that the reasoning underlying each participant's decision as to their confidence with one service over another would be multi-faceted, the flexibility of analytical method afforded by the type of protocol analysis used was considered most appropriate.

A subset of eight complete participant transcripts were initially divided amongst four individual researchers. Each researcher then used those transcripts to create an individual coding scheme that they felt effectively represented recurrent themes as they emerged within each transcript. The resulting four coding schemes were then collated to form a single coding scheme that was then applied to all 27 transcripts by two independent researchers. Finally, the resulting codings were collated once more, with any disagreements between the two final coders highlighted and discussed. While particular focus was afforded to themes / issues relating to genuineness, trustworthiness and security, coding was also completed for other factors of possible significance to the understanding the participants' backgrounds, perceptions and rationales.

Effects of physical linkage

Terms considered as being referent to *physical linkage* were frequently mentioned by the participants, and the majority of the participants felt that some aspect of the physical circumstances of the association between the artefact and the cafe was relevant to genuineness, security and trustworthiness for at least one of the conditions tested. These aspects fell into four general categories: *physical attachment*, *legitimacy*, *visibility* and *accessibility*. Each category is examined in turn.

Physical attachment referred to the degree to which an artefact was physically attached to some part of the cafe. This was frequently noted by participants. High degrees of physical attachment were often associated with the LCD screen: "*The screen looks bolted to the wall, physically bolted to the fabric of the building*", but less so with the leaflet, which was considered as just "*floating around*". Several participants observed that the poster, like the display, was physically fixed, although the lack of a securing frame (the poster was simply tacked to the wall) was commented on as something that could make the poster more susceptible to replacement or attack.

Legitimacy referred to the degree to which participants considered an artefact to be provided by the cafe. Terms used included "legal", "official" and "branded": "*once again it's branded and kinda looks consistent with everything else in the cafe*" and "*yeah . . . that looks official, that looks right*". However, the potential for forgery was also recognised: "*Well, if it wasn't the genuine one, it still could have said it was Bertorelli's couldn't it*"?

Visibility referred to the degree to which an artefact was in public view, and thus potentially less susceptible to an attacker being able to subvert that artefact undetected. The large LCD

display was rated highly in this respect: "*any kind of place wouldn't have screens and stuff that were interacting with computers and things without staff and management being very aware of it ... you couldn't scam that*". With regard to the poster, its relative lack of physical fixedness appeared as being made up for by its large physical size: "*The poster was more noticeable than the flyers because it's on the wall, and it's quite large, which makes it similar to the screen*". However, with the smaller leaflet, this effect was lost: "*[they] could be surreptitiously slipped in*".

Accessibility referred the to the degree to which participants considered the position of the artefact as being accessible by patrons. The leaflet was rated quite low in this respect, with comments such as "*anyone can go round and put in stuff*" frequently encountered.

Effects of virtual linkage

Every participant mentioned some aspect of the interaction between the laptop and the LCD screen that they encountered in the *synchronisation* and *interlock* conditions. Terms frequently encountered included "interaction", "information exchange", "transaction" and "dialogue". The mere presence of a degree of interactivity alone appeared to provide evidence for genuineness and trustworthiness: "*Because [the poster] doesn't provide any interactivity, I wouldn't trust it as much as the screen method*", and by increasing degrees of the complexity of interactivity, this effect was also increased: "*because it's a pretty complex interaction, you tend to think it was the real network*".

The specifics of the interactions encountered by the participant were however considered in different ways. The effect of *synchronicity* (agreement in time and content between the display and the laptop) for example was mentioned often as providing evidence of genuineness: "*cos it's changing quite quickly and it's staying exactly in sync... I think it shows that you are directly connected to the thing*". Conversely, with regard to the *Interlock* condition, comments more often made about *causality* (i.e. cause and effect within the interactivity). Sometimes participants referred to feedback from the display: "*I knew I was connected [to the right network] when I saw the results*", and one made specific mention of the security that they felt this provided "*[display-based protocols] feel more secure because there's more feedback*".

Finally, whether the data involved in the dialogue was personal to the participant was also an important factor for some: "*The message I chose was displayed [on the display], so I'm definitely connected to the right router*". And conversely: "*but [password on display] didn't even give me a chance to contribute myself ... so I'm not liking this one*".

Reasoning for the 'plateau' effect observed for measures of 'genuine', 'secure' and 'trustworthy'

The results of the experiment indicated that two sets of conditions were considered by participants as being equivalent in terms of genuineness, trustworthiness and security. These

were: *Password-on-poster / password-on-LCD display* (both password-based protocols), and *Synchronisation / Interlock* (both interactive protocols). Participant reasoning allowed for some means of describing how these effects came to occur. The password-based interactions and Synchronisation/Interlock protocols are discussed separately.

Password-based protocols: Only a small minority of participants made any attempt to directly compare the two forms of password-based protocols that were used in the experiment. In all cases they said that they were the same, but did not provide any clear reason why this was considered so. Many participants did however remark that the way passwords were used in the experiment was the opposite of how they normally expect to use passwords: in both cases the password was made public rather than private, and some (mis)interpreted the password as providing authentication of customers to the system, not the other way around (as was intended): "*so you've got to authenticate yourself to the machine*".

In terms of the previously discussed aspects of physical linkage, it is apparent that the differences between the forms of password-based protocol were considered as more similar than was expected. By deliberate design, the content of both poster and LCD display were made identical (i.e. equally legitimate), with the same orientations and of roughly equivalent physical size (thus equivalent in other aspects of physical linkage so far discussed - visibility and accessibility). However, in a purely physical sense, The LCD display was more firmly attached (physically fixed) to the wall than the poster and this was also deliberate (i.e., while a lockable frame for the poster could have been used, it was not).

Finally, it also seemed to some of the participants that the arrangement would be more secure in some sense if the password changed frequently. Some preferred the display to the poster for that reason: "*The beauty of the one on the screen is that they could change the password every 24/48 hours*". One theory was that this would make life harder for an individual attacker, who would, they reasoned, have to change the attack to accommodate new passwords.

***Synchronisation* and *Interlock* protocols:** Both protocols were considered highly novel, and there was a strong sense that the two protocols were unlike anything the participants had experienced before. Words such as "strange", "novel" and "weird" occurred frequently when participants described them. However, in terms of their value for genuineness, trustworthiness and security, participants appeared to be either unable or unwilling to discriminate between these two protocols. About half the participants made statements to the effect that the two protocols were similar in some sense. Reasoning for these statements fell into two general classes: *Intuitive response* and comments about *complexity.*

Intuitive response: Intuitive responses were vague on detail, reflected in comments such as "*I'm going back to these two... dunno why ... I just like them*", "*[these two conditions] make it harder to break in*" and "*[these two conditions] are ok, the rest no...*".

Complexity: Respondents frequently mentioned that both the *Synchronisation* and *Interlock* conditions were too complex to fake, but again were unable to discriminate between the two: "*[Interlock and Synchronisation] are the ones I completely trust.*", "*It would take a lot of resources to fake them.*" and "T*hey are different. They look like they are trying hard to make it secure. That's my impression*". However, in term of their security, the two pro-

tocols were considered as equivalent: "*I suppose [Interlock is] similar to [Synchronisation]. I don't think it's any more secure for being able to list a number of questions and a number of answers*".

Perception of risks relating to open-access wifi: Reasoning about the security value of the protocols

Some participants indicated that the use of public WiFi networks involves a trade-off, where the convenience of such a service is balanced against the activities for which it can be used. Others felt that it would be worthwhile trusting a public WiFi network for activities that they otherwise considered too great if other options were not available. For example, though some participants would not generally consider the use of open public WiFi to read their email, they indicated that they might if say, they were abroad and unable to access email any other way. The participants perceptions about security generally concerned the protection of their data: "*I think security is ... degrees of security to my vulnerability to some kind of crime really... access to my data ... [my] address book at one end to my debit card details at the other*". Eavesdropping relating to wireless communication was commonly mentioned as the most likely potential attack, and the need for encryption to protect their communications against such activity was frequently mentioned: "T*he fact that data I'm inputting into my computer that's going to their WiFi connection might potentially be being accessed by someone. I don't know enough about the technology to know more.*" and "*It feels less secure than if it was a wired Ethernet connection. I might be using their actual system, but it doesn't mean it's secure*".

5.6 Discussion and conclusions

5.6.1 The effect of physical / virtual linkage evidence upon user perceptions of situated service genuineness, trustworthiness and security

The primary hypothesis of the experiment was that user confidence in the genuineness of a service would be seen to increase if sufficient evidence was provided that linked that service to the immediate physical environment of the user. Within the experiment, two forms of linkage were considered and their effectiveness evaluated:

***Physical* linkage:** The linking of a digital service to the physical surrounding of the user by using artefacts that are more tightly attached to the immediate physical world (in a simple physical sense) than an intruder could contrive.

***Virtual* linkage:** The linking of a digital service to the physical surroundings of the user by using more interaction between physical artefacts and the digital service than an intruder could contrive.

The experimental results suggested that this hypothesis held initially for both forms of linkage. Participants rated a situated service as being more genuine, trustworthy and secure the stronger they were perceived as being physically or virtually linked to the cafe where the experiment was conducted.

However, the participants considered physical and virtual linkage as relative rather than absolute evidence for trustworthiness, and it was also clear that the participants thought about the roles of physical and virtual linkage in a variety of ways. The perceived difficulty of faking each physical / virtual artefact was a key factor in this regard, but in very few cases were the participants estimating the feasibility of fakery in a technical sense; rather, they appealed to factors that were familiar to them. For example, the creation of leaflets was taken to be easier and less expensive than the printing and mounting a poster.

Within the more technical protocols examined (*Synchronisation* and *Interlock*) participant reasoning about the potential for fakery invoked similarly interesting ideas. With the *Synchronisation* protocol, the difficulty in subverting the system was based upon the difficulty of getting the timing just right, whereas in the interlock protocol it was the the difficulty involved in the attacker guessing the chosen phrase.

5.6.2 The provision of a secure and usable protocol to support secure device association in a situated-service usage scenario

Within the experiment, a number of simple password-based protocols were compared with two substantially more complex and interactive methods of authentication. In reality, none of the password-based protocols had any actual value for authenticating the network as being truly secure. However, only a few participants made note of this fact: "*you're just typing in something, it could be another access point that accepts the exact same password... It doesn't mean that it's Bertorelli's*".

Of the remaining two protocols examined, while the *Synchronisation* protocol was able to offer some protection against an *evil-twin* attack, it offered little to no protection against the *man-in-the-middle* attack. Thus, in actuality, only the remaining protocol, *Interlock*, could be considered truly secure as it was the only protocol to offer actual security value against the threats of both *man-in-the-middle* and *evil-twin*. However, participants tended to consider the security and trustworthiness value of both protocols to be equivalent.

Finally, the usability of both protocols was a frequent topic of discussion, with the most common themes being the time and effort required to complete the authentication process and concerns about the use of publicly visible screens.

Time and effort: Both *Synchronisation* and *Interlock* protocols took substantially more time and effort to complete than did the password based protocols. This was particularly the case with *Interlock*, where a multiple stage interaction was involved in order to complete the authentication process. While several participants found the *Synchronisation* and *Interlock* protocols to be engaging, and even fun, the procedure involved in the Interlock protocol was hard to explain to the non-technical. Consequently, it was frequently described by

participants as being complicated and confusing. Broadly speaking however, ease was taken to be indicative of dubiousness, and difficulty was taken to be evidence for genuineness. Thus while future revisions to the design of the Interlock protocol could easily address the concerns raised, designers should be mindful of the degree to which they make their solutions simple to use!

Use of publicly visible screens: The use of a publicly visible display was considered less than ideal, and participants were concerned with the ease by which they could be observed by others as they interacted with it. Further usability issues also derived from the use of a public display which may not be clearly visible at a distance (making comparisons between near and far screens - i.e. the user's laptop and the cafe's display - potentially difficult), or well suited to a real-life situation where multiple users may need to interact with it simultaneously.

5.7 Chapter summary

In this chapter, a scenario was created whereby several co-present situated services were made simultaneously available in a real-world cafe environment, thus more closely reflecting how real-world pervasive computing services might come to be encountered. Within this scenario, a semi-laboratory experiment was conducted in which participants were invited to assess a number of wireless Internet provision services (WiFi 'hotspots'), in terms of the degree to which they considered them as being trustworthy, genuine and secure.

To assist participants in this process, the experiment sought to utilise and investigate how certain *location-based* artefacts in the physical and virtual worlds could be leveraged in order to offer the participants evidence that one wireless situated service - made available amongst a plethora of others - was genuinely provided by the owners of the cafe. Further, the experiment sought to offer methods and protocols whereby connections to trustworthy services could be established securely, thus addressing a specific characteristic and problem of pervasive computing usage: the need for *secure ad-hoc device association*.

While the experiment did not measure actual trust investment behaviour, it did provide some means of understanding the conditions in which a real user of a typical situated service might come to invest their trust. In performing this investigation therefore, the thesis was able to develop further the findings of the previous chapter and contribute a deeper understanding of **RQ's 2, 3 & 4**. Participants were found to be security conscious and wary of public wireless services, and this wariness was only increased by having multiple wireless services in the same physical place. Most participants, as was observed in the questionnaire deployed in the previous chapter, had little in the way of past experience with public access ICT services.

Participants in the study demanded substantial reassurances that the services they chose to use were genuine, secure and trustworthy and this was reflected in their reticence to engage in risky activities (checking email, making a payment etc) using wireless services that did not offer such assurances. That their confidence in the trustworthiness and security of a wireless

service was however found to increase with the strength of both physical- and virtually based linkage evidence provided some evidence that the use of linkage as an evidential cue was an effective technique. Supporting this effect further, in the case of services that included virtually-linked evidence, though participants found it hard to compare two unfamiliar and relatively complex protocols, overall, they did find them trustworthy enough to suggest that they might use such services for more risky activities, including financial transactions.

In the next chapter, the concept of digital- to physical-world linkage as means of providing evidence of service trustworthiness would be explored further with an appreciation of the effect of a final type of artefact: *Other people*.

Chapter 6

The effect of other people: *Social linkage* evidence as a means of increasing user perceptions of situated service trustworthiness

6.1 Chapter overview

The physical world, as Kindberg et al (2002) [72] observed, can be considered as being made up of three general categories of entity: *people*, *places* and *things*. In the previous chapter, an experiment was conducted whereby various artefacts (things) were made available in the physical world (place) that were then leveraged to provide evidence as to the genuineness (and by extension) trustworthiness of a wireless situated service (a WiFi 'hotspot'). As constituting that evidence, two forms of linkage, by which the wireless situated service could be linked with the physical world in which it was encountered, were investigated: *physical* and *virtual* linkage. The results of that experiment showed that, by increasing the degree to which people consider evidence as being physically linked, there was a corresponding effect of increasing their perceptions of the genuineness, trustworthiness and security of the service to which that evidence was associated. Further, through the introduction of evidence that was virtually linked, this effect was increased further than physically linked evidence alone.

In this chapter, an experiment was developed to examine the use of a possible third linkage mechanism that is based upon the remaining category of entity as observed by Kindberg et al: *people*. This final form of linkage is termed '*social linkage*', and evidence of its existence would be provided to users through extending a typical wireless network service discovery

protocol.

6.2 Introduction

Urban pervasive computing describes a world where digital and physical worlds co-exist within the same space. Within such an environment, numerous digital services and physical people are able to encounter, associate and interact with one another freely and wirelessly. The wireless digital world is, however, not directly perceptible to humans. Thus, to become aware of the digital world around them, it is necessary for people to use some form of computing device (such as a mobile computer with wireless networking capabilities) that has the ability to both detect and communicate with wirelessly delivered digital services. Thorough the use of such devices, physical people are able to 'see' and communicate on an ad-hoc basis with digital services and vice-versa.

Within both worlds, many individual entities can co-exist within a given shared space. However, communication between the two worlds is (by and large) performed one-to-one; that is to say that a given digital service present in a particular physical place does not typically share its knowledge of the physical world with other digital services within the same space. Likewise, a given physical person does not generally share their knowledge of the digital world with other physical people (unless questioned directly). As discussed in previous chapters, this poses a problem from the perspective of a person uncertain over which of a number of potential digital services it is safe to connect to. Should a person wish to take advantage of one particular wireless situated service that is available alongside a plethora of others, they must accept a degree of risk that the service that they choose may not be bona-fide. Further, they must also accept that the act of merely making a connection to any service may be enough to compromise the security of their data.

This particular problem, that users are not afforded information as to the potential trustworthiness of a given wireless situated service prior to connection with that service, forms the primary focus of this chapter. Current wireless service discovery protocols, such as those bundled as part of the two most used commercial computer operating systems (Windows 7, Mac OS X 10.6[1]) are a case in point. Typically (i.e. without additionally installed software or substantial technical knowledge), the service discovery tools offered by the major operating systems offer users only a limited amount of information about the services that they detect. Generally, this information is limited to the name of a detected service, whether or not that service offers encrypted communication and perhaps an approximated measure of its radio signal strength. Little, if any of this information is able to provide evidence to the user that any detected service can be considered as being trustworthy. Service names can be copied or otherwise subverted, signal strength is subject to various uncontrollable physical phenomena and communicating with a bogus service via an encrypted communication channel is perfectly feasible, if not entirely straightforward for an attacker to arrange.

[1] Current as of writing in late 2010.

6.2.1 People as evidential cues to digital service trustworthiness: *Social Linkage*

Social linkage is defined here as "*the utilization of physical people present within a bounded space to create a link between a wireless digital service and the physical space in which that service is made available*". Within this definition, the use of the term *bounded* has been used to illustrate the point that that the digital and physical worlds share some common features that are important. Physical people are often contained within particular areas of physical space by structures through which they cannot pass (e.g. walls). Similarly, if to a lesser extent, digital radio signals are limited in terms of their range by measure of their signal generating power and the effect of signal attenuation caused by surrounding physical structures. To a degree therefore, in a given space, both physical people and wireless digital services are bounded somewhat by the physical configuration of that space; a wireless digital service is of limited range, and potential users of such a service must themselves be physically present within that range. Indeed, it is the specific type of digital services that exist within these parameters ('*situated*' services) that this thesis is primarily interested (see section 1.3).

Regarding the use of situated services such as public access WiFi 'hotspots', as such a service is typically offered as a value-added service to paying customers of the host venue, the range of that hotspot is assumed and generally intended to be limited to the walls of the host venue (though various radio transmission phenomena will often expand this range somewhat further than might be intended). As the range of the service of a WiFi 'hotspot' is limited, it could be considered reasonable to assume that the source of the signal is to be found somewhere within the confines of the venue. Thus, the assumption might easily be made that the infrastructure required to provide an encountered WiFi 'hotspot' is likely to be owned and administered by the proprietors of the venue in which it is found. In any given space however, several independent WiFi 'hotspots' can, and often do co-exist. As demonstrated in the previous experiment (chapter 4), the creation of such services can be achieved through little more than a WiFi-capable laptop computer and basic knowledge of networking.

In the social world, people often seek and make use of social information that is embedded in their social networks in order to reduce uncertainty about their choices and actions (Mohtashemi & Mui, 2003 [105]). This effect is made particularly salient in situations that contain aspects of ambiguity, such as might be found in a situation where a number of situated services - all purporting as being genuine - are presented to a naive user from which to choose. In the absence of prior interaction history, that others are seen to have used, or are using a service has been found to increase an individual's personal trust in that service (e.g. Boyd, 2002 [16]). Creating a sense of user community, i.e. that others have, and continue to use a particular service is a technique used widely on risk-relevant services that are made available on the WWW. User-generated and publicly visible vendor / product reviews are, for example, a common feature of e-commerce websites such as *eBay* and *Amazon*.

In the pervasive computing world, individual people are linked to the digital services that they are using by extension of the individual networkable devices that they carry on their person. Thus, if information could be provided to the user that indicates the number of digital devices that are connected to each wireless digital service that is detectable in a given physical space, we might reasonably conclude that they could relate that figure to

the number of physical people present within that space. This potential is the basis for the proposed *social linkage* channel. *Social linkage* is a bond created between physical people in a physical space and a wireless digital service in digital space. The strength of this bond can be measured as a product of the proportion of total possible physical service users (i.e. the number of physical persons present who could conceivably be using that service) relative to each of the digital services locally present (as is detected by a user's laptop). Through providing information about this relationship to naive users via the network connection software of their electronic device, we suggest that their perception as to the trustworthiness of one service over another would be positively affected.

6.3 The design of an experiment to examine the effect *social linkage* upon user perceptions of situated service trustworthiness

6.3.1 Research questions and hypotheses

The experiment sought to develop and evaluate methods by which a basic service discovery protocol (WiFi network service detection) could be extended to include information as to the current usage of each service that it currently detected as being within range. The additional information that the protocol would provide to users would be a representation of the number of people currently using any specific local wireless service. As WiFi range is relatively short, users would be feasibly be able to validate this information further by comparing the system generated information with the numbers of physical people (who could plausibly be using such a service) who were nearby.

We predicted that, within a situation where multiple services were detected as being available, and in the knowledge that only one of those services was definitely genuine, users would utilise information as to the current usage levels of the services detected to assist them in their choice of service with which to connect.

Thus, the first hypothesis (**H1**) of the experiment was: *As the proportion of virtual users shown as being connected to one wireless network increases relative to other networks that are available, user confidence in the genuineness of that network will increase.*

The term *virtual users* refers to the number of people (asserted by the protocol) that are actively connected to a particular wireless digital / situated service at that time. Conversely, *physical users* refers to the number of people in a space that are observable as using digital devices that could plausibly be using any wireless digital service. That these two figures could be compared, and thus potentially affect a sense of social-linkage, a second hypothesis (**H2**) was also formed: *For a given network, as the number of users reported by the system increases or decreases relative to the number of plausible* physical *users (i.e. that the user can physically see in their immediate environment), confidence in the genuineness of that network will reduce.*

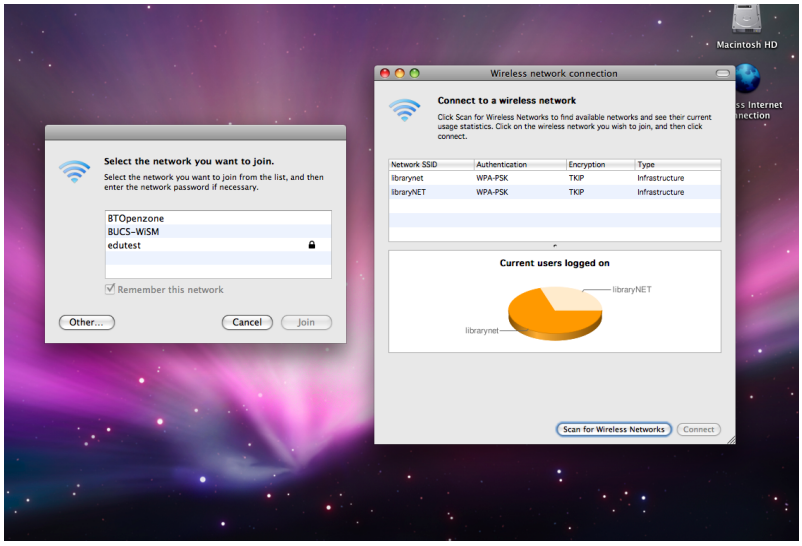


Figure 6.1: *Social Linkage* Wireless service connection UI (right). Look and feel mirrors Macintosh OSX 10.5 operating system. Actual Mac OS wireless service connection software (left) is shown alongside for comparison.

6.3.2 The design of a service-discovery protocol to support *social linkage*

As with the experiment conducted in the previous chapter, a new protocol would need to be developed that would be used at the point at which the user, having ascertained the presence of more than one situated service, is in the process of deciding which of them is most likely to be genuine. *Genuine* in this sense can be understood as meaning '*made available / administered by the management of the host venue and of benevolent intention*'. Typically, wireless digital service discovery and connection is facilitated using software embedded within the user's operating system (OS). To create a protocol that would support evidence for *social linkage*, a new wireless service discovery procedure and user-interface was developed by the researcher to replace the one that was provided by the operating system (figure 6.1). OS wireless network connection software typically offer some combination of the following information for any service that it encounters: service name (SSID[2]), levels of access restriction if present (password required etc) and a measurement of radio signal strength. The *social linkage* protocol that was devised simply extended this information by making available socially derived information about current user activity – i.e. a representation of the number of devices currently connected to each detectable service.

The interface designed for the experiment was carefully created such that it would appear to function as an integrated part of the OS. Thus, substantial attention was paid to the aesthetics and functioning of the UI so as to closely match the look-and-feel of the host OS (Apple OS X, v10.5). However, while purporting to be able to both locate and communicate with actual wireless networks, this process was in truth simulated and thus a deceit.

[2]SSID: Service Set IDentifier.

Figure 6.2: Evaluation venue: University of Bath Library social space.

6.4 Experimental design

6.4.1 Method

As with the experiment reported in the previous chapter, evaluation of the *social linkage* protocol was conducted within a semi-field setting. *Semi-field* referred to the fact that the study was conducted in a public place, but that the experimenter was present and the participant aware that they were undertaking a controlled experiment. As a semi-field experiment, the choice of venue used in the experiment required careful consideration. The experiment involved the use of real people as an experimental device, and thus needed to be both publicly accessible yet to also have a number of persons (who would not know that they were involved in the experiment) to be both visible and using Internet-capable mobile devices (e.g. laptop computers).

The choice of venue used for the experiment was the ground floor social space within the *University of Bath Library* (figure 6.2). This space was chosen as it attracted a large congregation of people who would use the space to work and socialise throughout any given day. As the space offered its own wireless Internet service[3], a good proportion of visitors to the space were generally to be found using laptop computers.

The use of deception required by the instantiation of the protocol, and the utilisation of members of the public as an experimental device were considered in terms of their ethical impact on participants. Ethical approval for the study was obtained from the University of Bath department of Psychology.

[3]To reduce possible confounds relating to previous experience with the existing wireless network service ('BUCS Wireless'), this service name was not used in the experiment.

6.4.2 The independent variables and the creation of experimental materials

The experiment was a within-subjects design with two independent variables (IVs): *distribution of users associated with each service*, which related to **H1** and *accuracy of numbers reported* which related to **H2**.

To support IV1, *distribution of users associated with each service*, two situated services (both WiFi hotspots) were created, with the names (SSID's) *Librarynet* and *LibraryNET* respectively. Participants were informed that the names of the services were not to be considered in their reasoning, save for the consideration that the two names used were plausible variants of one actual genuine service. Both networks were reported by the system as being encrypted using WPA-PSK[4]

IV2, *accuracy of numbers reported*, referred to the manipulation of the true number of physical people present who could plausibly be using wireless services. The purpose of this manipulation was to encourage the participant to consider the number of people who were actually present in the experimental venue (i.e. the strength of social linkage). To support IV2, two figures were obtained prior to each evaluation trial through the researcher conducting a headcount of people present in the venue. These were: 1) total persons present and 2) total number of those persons currently using laptop computers. For the purposes of the experiment, it was assumed that all laptops that were in use within the venue were capable of connecting to the Internet wirelessly.

To form the *social linkage* based evidential cue, two simulated WiFi services were presented to participants via the *social linkage* UI, each with an associated figure that represented the number of devices that were currently connected to them. The form of the representation was a simple pie chart (figure 6.3).

Two sets of conditions were created to evaluate the two hypotheses / IVs of the experiment. To test hypothesis / IV one, only the relative proportions of users associated with each of the two services was shown. To achieve this, the total number of laptop users present (as ascertained by the researcher prior to each evaluation) were divided across the two bogus networks to form eight configurations / conditions, shown below in table 6.1.

Hypothesis two stated that reducing the accuracy of numbers reported in the system would have a corresponding reduction effect upon user confidence. To evaluate hypothesis / IV two, the proportional distribution of laptop users was fixed at the level to which participants were most confident that the service they chose was genuine (i.e. from the IV1 trials). For the conditions presented for hypothesis two, the actual number of users connected to each of the two services was made available on the pie chart shown by the *social linkage* UI. Seven conditions for hypothesis two were generated by manipulating the range between the numbers of users reported by the UI and the true number of plausible users present (as ascertained by the headcount of actual laptop users). The resulting conditions are presented in table 6.2.

[4]WPA-PSK: WiFi Protect Access - Pre-Shared Key. This variation of WiFi communication encryption is typically found on home wireless networking services and is often enabled by default.

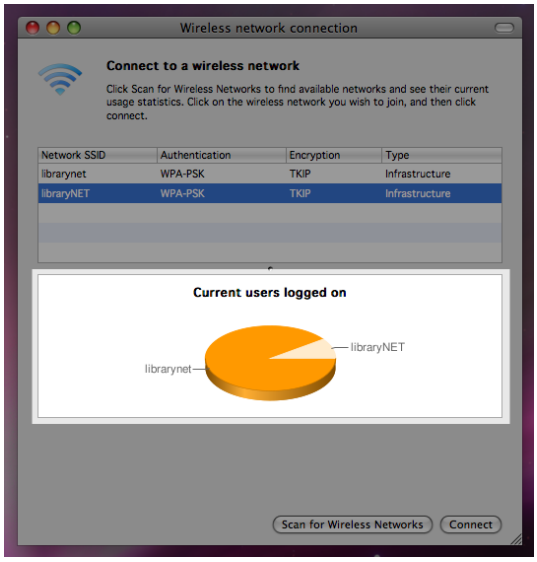


Figure 6.3: *Social Linkage* Wireless service connection UI showing two discovered services. Representation of proportion of users currently associated with each service is highlighted.

Condition	**% Virtual users assigned to librarynet**	**% Virtual users assigned to libraryNET**
1	10.0%	90.0%
2	20.0%	80.0%
3	30.0%	70.0%
4	40.0%	60.0%
5	60.0%	40.0%
6	70.0%	30.0%
7	80.0%	20.0%
8	90.0%	10.0%

Table 6.1: Conditions for IV1: *distribution of users associated with each service*

6.4.3 Dependent variables and experimental measures

Two dependent variables (DVs) formed the experiment measures. These were *confidence in the genuineness of the service chosen* and *confidence in the genuineness of the service* ***not*** *chosen*. Quantitative measures of both forms of confidence were obtained using an interactive slider mechanism that provided a 500-point scale from 0 [not at all confident] to 500 [very confident]. To gain additional qualitative data relating to participant reasoning as to the two DV's, participants were also encouraged to "think-aloud" as they performed the tasks set. Participant vocalisations were captured using a dictaphone and were supplemented by notes taken by the researcher. Finally, as the number of people present in the space varied substantially across different times of the day and different days of the week, records for both visitor numbers and laptop users were recorded prior to each experimental trial and were included for use in subsequent analysis as co-variates.

Condition	Accuracy of numbers reported (as derived as a % of the true total no. of laptop users present)
1	25.0%
2	50.0%
3	75.0%
4	100.0%
5	125.0%
6	150.0%
7	175.0%

Table 6.2: Conditions for IV2: *accuracy of numbers reported*

6.4.4 Experimental procedure

Prior to commencement of each trial, the experimenter obtained the total number of 1) people present in the experimental venue and 2) the number of those people who were currently using laptop computers. If the total number of people present was less than 30, or the number of laptop users was less than half the total number of people present, the trial was abandoned. Both numbers were entered into the system that supported the *social linkage* UI.

Each experimental trial was conducted in two phases. Phase one was the evaluation of conditions relating to **H1**, during which participants were exposed to eight different configurations of active user distribution across the two WiFi services. Phase two was the evaluation of conditions relating to **H2**, during which participants were exposed to seven variations of the actual numbers of users associated with each service. The duration of each complete experimental trial was between 25-30 minutes.

Each participant was recruited outside of the study venue, where consent was obtained in writing and brief instructions as to the nature of the study given (see appendix C, sections C.1 and C.2). The introduction provided to the participants is presented thus:

> *Imagine you are visiting a public library to try to get an Internet connection. You have your laptop computer with you, and you know that this library happens to have a wireless Internet access point that you can use during your visit.*
>
> *When you sit down at a table and try to connect to the wireless network you see that there are in fact two different wireless networks that both appear to be provided by the library.* ***You know that the library genuinely provides one of these networks. The other network may be from another source pretending to be the library and should be considered as potentially fake****. You are aware that wireless networks are vulnerable to attack; somebody might have created an entirely fake network to which you can still connect, potentially giving away passwords or other information and inadvertently giving the faker access to your computer.*

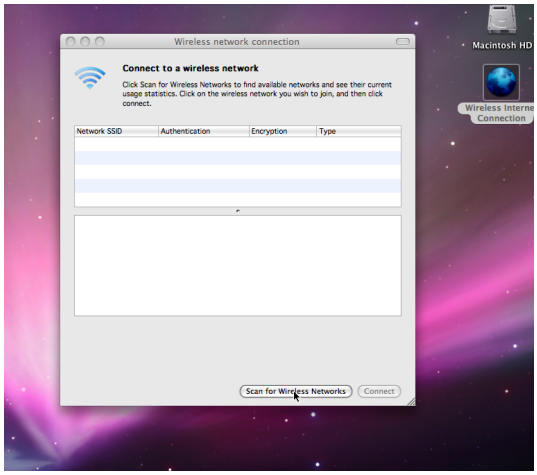


Figure 6.4: *Social Linkage* UI: Default state, prior to discovering available wireless network services.

The participant was then guided through the venue and seated at a table that had a laptop computer made available for their use. Instructions for discovering and connecting to currently available wireless networks using the *social linkage* UI were offered by the researcher. Each participant was told to assume that all of the data that they would encounter during the experiment was to be considered as having been collected in real-time.

When the participant felt ready to begin, they were instructed the use the *social Linkage* UI[5] to 'scan for wireless networks' (figure 6.4).

Throughout the experiment, the researcher remained seated next to the participant. The participant completed all experimental trials at their own pace, with eight trials presented for **H1** (in randomised order), followed by seven trials for **H2** (also randomised). As the data required for the generation of the conditions in **H2** were dependent upon the results gathered for **H1**, counterbalancing those two phases of the evaluation was not possible.

For each of the 15 trials, the participant was presented with the proportion and / or number of users currently associated with the two fictitious wireless services *librarynet* and *libraryNET* (figure 6.5). To complete each trial, participants were asked to select which of the two services they considered to be most genuine, and then to connect to that services through clicking a button labelled 'connect'.

Upon choosing the service that they considered as being most likely to be genuine, each participant was then asked to rate each of the two services (i.e. the service that they chose, and the service that they *did not* choose) in terms of the question '*how confident are you that the network you have chosen is [not] a genuine wireless hotspot offered by this venue?*'. Participant measures of confidence were collected using a sliding scale of *not at all confident* to *very confident* that could be manipulated using a mouse connected to the laptop (figure 6.6).

[5]The name of the UI as seen by the participants was given simply as 'Wireless Internet Connection'.

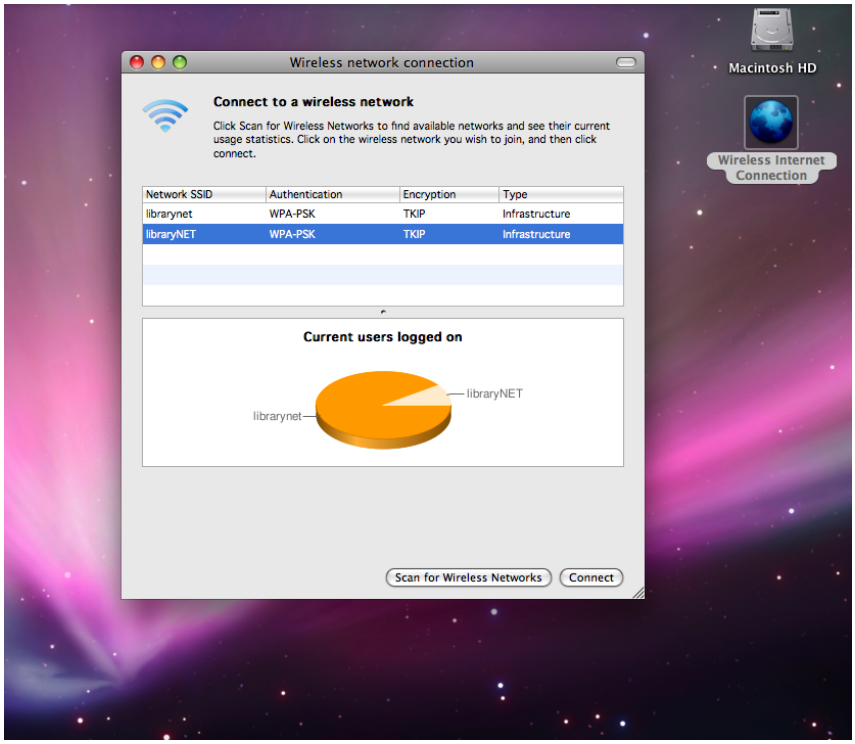


Figure 6.5: *Social linkage* UI: Two available wireless network services found. The distribution of users associated with each service is presented to the participant as a pie chart.

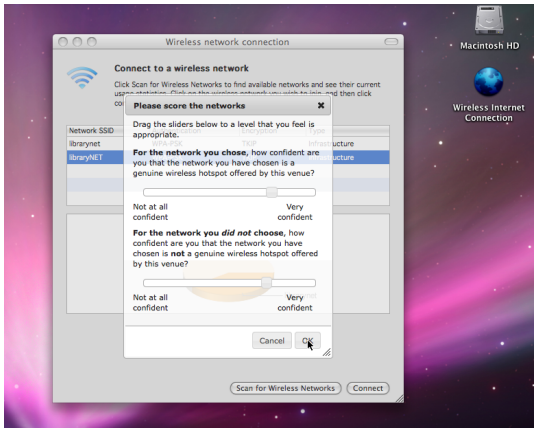


Figure 6.6: *Social linkage* UI: Data collection, confidence scores for the network services *chosen* and *not-chosen*.

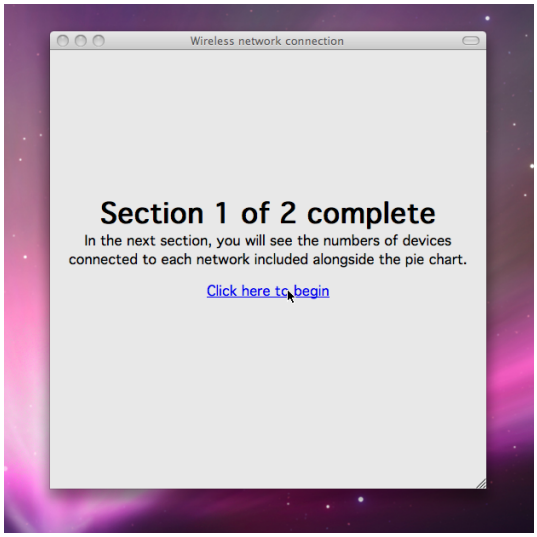


Figure 6.7: Evaluation UI: Separation of conditions relating to hypothesis one and two.

Conditions for hypothesis one (only proportional distributions of users shown) and hypothesis two (both proportions and actual numbers shown) were separated with the screen presented in figure 6.7.

Upon completion of all 15 experimental trials, the participant was finally asked to answer two questions:

- How many people did they think were actively using [any] wireless Internet service in this space?
- How did they come to obtain that number?

The participant was then debriefed and any questions raised during the experiment were answered.

6.4.5 Participants

The participants were 24 individuals (Gender split: M=13 (54%), F=11 (46%)), recruited by email invitation and general opportunity sampling. 20 participants (83%) were current University students (undergraduate / postgraduate) and the remaining 4 (17%) were members of University staff. The modal age range of participants was 18-25.

6.5 Results

6.5.1 Effects of proportional distribution of users associated with librarynet and libraryNET

Mean confidence scores for the service chosen as being the genuine service, and the service *not* chosen as being the genuine service were calculated across the eight conditions of IV1 (*distribution of users associated with each service*) and are presented as a table (table 6.3).

Condition	**% Virtual users using *librarynet***	**% chose librarynet as genuine service**	**Mean confidence score for chosen service**	**Mean confidence score for non-chosen service**
1	10.0%	1 (4.2%)	362.21	310.04
2	20.0%	1 (4.2%)	273.71	192.25
3	30.0%	1 (4.2%)	193.21	146.83
4	40.0%	7 (29.2%)	106.96	93.25
5	50.0%	21 (87.5%)	90.54	79.71
6	60.0%	24 (100.0%)	189.08	139.67
7	70.0%	23 (95.8%)	236.00	188.46
8	80.0%	24 (100.0%)	352.25	282.08

Table 6.3: Mean confidence rating scores for the service chosen / not-chosen for the eight conditions of IV1 (*distribution of users associated with each service*).

Confidence scores for the service *chosen* as being the genuine service

Mean confidence scores (range 0-500) for the service that was chosen as being genuine were analysed across all eight conditions of IV1 *distribution of users associated with each service* using a one-way repeated-measures Analysis of Covariance (ANCOVA). To control for possible effects due to variance in numbers of people / laptop users between participants, the values for *total number of people present* and *total number of laptop users* were included in the analysis as co-variates.

A significant main effect of *distribution of users associated with each service* [Wilks' Lambda = .35, F(7,15) = 3.930, p=0.01] was observed, and a multivariate partial eta squared value of .64 suggested a large effect size. No significant co-variate effect was found for *total number of people present* [Wilks' Lambda = .81, F(7,15) = 0.487, p=0.83 n.s] or *total number of laptop users* [Wilks' Lambda = .76, F(7,15) = 0.693, p=0.68 n.s] indicating that the main effect was not influenced by variances in the number of people / laptop users that were present in each trial.

Pairwise comparisons (repeated measures t-tests) were then performed on the scores between

successive condition pairs in which the distribution of users associated with each of the two services is skewed first towards service one (librarynet) and then to service two (libraryNET). Six of the seven paired comparisons were found to be significant at the 0.05 level, indicating that, as the proportion of virtual users assigned to a network fell (in intervals of 10%), participant confidence in their eventual choice of that service (i.e. as being genuine) also dropped significantly. However, as no significant difference in confidence score was observed between conditions 4 & 5 [t(23) = 1.01, p=0.32 n.s], it also appeared that this effect had a threshold. When the proportional split approached 60/40, participants considered both the librarynet and libraryNET services to be equally (un)likely as being genuine.

Confidence scores for the service chosen as *not* being the genuine service

Mean confidence scores for the service that was chosen as being *not genuine* were analysed across all eight conditions of IV1 *distribution of users associated with each service* using a one-way repeated-measures Analysis of Covariance (ANCOVA), with the same co-variates values: *total number of people present* and *total number of laptop users*.

No significant main effect was observed for *distribution of users associated with each service* [Wilks' Lambda = .49, F(7,15) = 2.194, p=0.09 n.s], and no significant co-variate effect was found for either *total number of people present* [Wilks' Lambda = .66, F(7,15) = 1.129, p=0.40 n.s] or *total number of laptop users* [Wilks' Lambda = .77, F(7,15) = 0.639, p=0.72 n.s]. Though the ANCOVA analysis approached significance for the main effect, it appeared that participants did not consider the differences in proportional distribution in terms of their confidence about the service that they rejected. Examination of the percentage mean confidence scores showed that confidence in the service not chosen was 36% and no score rose above 65%.

6.5.2 Effects of the accuracy of numbers reported by the *social linkage* UI

Mean confidence scores for the service chosen as being the genuine service, and the service *not* chosen as being the genuine service were calculated across the seven conditions of IV2 (*accuracy of numbers reported*) and are presented below as a table (table 6.4).

Confidence scores for the service chosen as being the genuine service

Mean confidence scores for the service that was chosen as being genuine were analysed across all conditions for IV2 *accuracy of numbers reported* using a one-way repeated-measures Analysis of Covariance (ANCOVA). As with previous analyses, values for *total number of people present* and *total number of laptop users* were included in the analysis as co-variates.

No significant main effect was found for *accuracy of numbers reported* [Wilks' Lambda = .65,

Condition	System reported accuracy (as % of actual count of laptop users)	% chose librarynet as genuine service	Mean confidence score for chosen service	Mean confidence for non-chosen service
1	20.0%	24 (100.0%)	363.96	381.79
2	50.0%	24 (100.0%)	349.21	336.33
3	75.0%	24 (100.0%)	396.38	375.21
4	100.0%	24 (100.0%)	394.21	351.33
5	125.0%	24 (100.0%)	398.58	375.83
6	150.0%	24 (100.0%)	399.67	354.25
7	175.0%	24 (100.0%)	409.75	357.54

Table 6.4: Mean confidence rating scores for the seven conditions of IV2 (*accuracy of numbers reported*).

F(6,16) = 1.443, p=0.26 n.s], and no significant effect was observed for either co-variate. Both results indicated that participants did not consider the accuracy of the numbers of virtual users (as presented to them by the *social linkage UI*) as being relevant to their confidence about the genuineness of the service they chose.

Confidence scores for the service chosen as *not* being the genuine service

Analysis of the mean confidence scores for the service chosen as being *not* genuine using ANCOVA found no significant main effect of *accuracy of numbers reported* [Wilks' Lambda = .68, F(6,16) = 1.242, p=0.34 n.s]. No significant effect was observed for either co-variate. Again, both results indicated that participants did not consider the accuracy of the numbers of virtual users (as presented to them by the *social linkage UI*) as being relevant to their confidence about the genuineness of the service they rejected.

6.5.3 Participant estimates of the numbers of people actively using a wireless Internet connection

Mean counts of laptops (actual) and laptops (as perceived by the participants) were calculated and are presented as a table in figure 6.5 below.

On average, participants underestimated the true number of laptop users by around 15%. When probed as to the method by what means this figure was obtained, 71% of participants stated as basing their estimate on guesswork "*more than 20, probably more like 30*", "*I'd guess about 30 in this room*". While several participants used the simple strategy of counting laptop users in some proportion of the room and then multiplying it up: "*at least 25 in sight, times by two for the rest*", most reported their figure as being based on an immediate

Mean no. of laptop users (actual)	Mean no. of laptop users (perceived)	% Diff
41	35	-14.60%

Table 6.5: Comparison of (mean) numbers of laptop users as ascertained by the researcher vs. (mean) numbers of laptop users as reported by participants.

appraisal and / or intuition. Only seven participants (29%) actually made direct efforts to take an accurate count of the actual number of laptop users in the room.

6.5.4 Participant reasoning about the impact of *social linkage* upon perceptions of service genuineness and trustworthiness

IV1 conditions (*distribution of users associated with each service*)

Within the eight IV1 conditions, only the relative proportions of virtual users assigned to librarynet and libraryNET were displayed. Participants - while generally appreciating its purpose - were confused as to why the system did not display the actual numbers of users that were associated with each service: "*If [the system] knows the proportions, it must know how many people are using librarynet*" and "*it's not much use if [the interface] doesn't show numbers*". Despite these reservations however, in conditions where the relative proportion of virtual users assigned to one service over another was high (i.e. more than 60% in favour), participants did use the pie-chart visualisation as guidance in their decision making. To do this, they typically many stated as favouring a 'herd mentality' type heuristic: "*I'll go with that..Cos More people are using that one*". though confidence in their selection appeared somewhat vague: [Researcher] "*Does having more people make it more genuine?*" [participant:] "*I dunno, I guess so*". However, as the relative proportion of virtual users assigned to each service approached equal (60% or less), participant confidence scoring bottomed out. Several participants stated that, had the experimental design permitted them to do so, they would have rejected both services: "*They're both pretty much the same. If I knew one was definitely fake I wouldn't want to connect to either... Because I have to pick one, i'll pick this one at random*" and "*I would be suspicious of both [librarynet and libraryNET]*".

IV2 conditions (*accuracy of numbers reported by the social linkage UI*)

The display of numbers alongside the proportional distribution was positively received by the vast majority of participants, many of whom which had previously made comment that they should have been included in the first place. Though no data exists to validate it, there was a general sense that the decision making process was shortened substantially by their inclusion. This was certainly the case in some trials, where only one virtual user was

assigned to one of the services: "*That's definitely the fake one.. well, there's only one person on it*".

When the numbers of virtual users connected to the minority service were less than five, participants considered the service to be similarly likely to be fraudulent. However, when numbers advanced above five, a small number of participants remarked that their confidence in the majority service was compromised: "*I don't know about this now. I'm beginning to wonder if the one with most people on it is the wrong one... There's too many people on the other one, if it were fake they would know*". The instructions supplied to the participants advised them to consider that the numbers reported to them by the *social linkage UI* were to be considered as being 'live'. As the numbers reported changed between trials, participants rationalised this change in a number of interesting ways. Generally, the participants considered the change as either being due to people coming or going that they could not see, or as connections to the two services being linked to the use (or disuse) of Internet browsing software.

Use of laptop / browser as being indicative of connection to Internet: Within the design of the experiment, an assumption was made that all laptops present within the experimental space were connected to a wireless Internet service. Of course, in reality this is not necessarily the case. A small number of participants stated this as part of their reasoning: "*Just because they are using a laptop doesn't mean they are on the Internet*". However several participants made the incorrect assumption that sustained connection to a WiFi service was dependent upon the activation of a computers Internet browsing software: "*If they're not using [their Internet browser], they're not connected to the Internet*". Thus, when the numbers reported by the *social linkage UI* changed between trials, their reasoning reflected this: "*They must not be using the Internet...Maybe they shut [their Internet browser] down*". One participant was particularly convinced by this, stating their own experience: "*On my laptop I have to reconnect to wireless when I turn Internet Explorer off*".

No single participant made mention of the fact that the relative proportion of virtual users assigned to *librarynet* and *libraryNET* did not change, despite the actual numbers being assigned to them changing with every trial. While it is true that the distribution of numbers shown did reflect the correct proportional distribution shown, it is surprising that no-one seemed to find the situation unusual.

Visibility of people: Most participants stated as having experience of wireless internet connection, both at home and elsewhere. As such, though rarely discussing the phenomenon on any technical basis, they appeared aware that the radio signals used by WiFi could feasibly be detectable outside of their immediate physical location. This was reflected directly in several comments: "*I can't see all the people the computer can see*" and "*[the system] must be detecting other people... [people] upstairs maybe*". When asked how this judgement came to be, the accuracy of the social linkage UI was never questioned. Instead, participants made reference to the physical laptop users that they could see: "*no-one has moved, so [the system] must be picking up people somewhere else*". Generally speaking however, participants appeared to be following the same 'herd mentality' logic as observed in the IV1 trials.

6.6 Discussion and conclusions

In this experiment, the presentation of socially-derived *social linkage* evidence was considered in terms of its effect upon user perceptions of situated service genuineness. The results of the social linkage experiment gave rise to two general findings: 1) evidence that the inclusion of socially-linked evidence was considered as being useful in a situated service usage scenario where several similar services were found to be co-present, but 2) that the evidence provided, though considered in terms of the number of plausible service users physically present was (for a number of good reasons) not considered as being convincing enough to warrant evaluation above a superficial level.

The primary hypothesis of the experiment was that, as the proportion of virtual users apparently connected to one wireless network increased relative to other networks available, user confidence in the genuineness of that network would also increase. The results of the experiment tended to support this hypothesis, providing that, given the two-potential-network scenario examined, the proportion of users assigned to one network was more than 60%. Most participants considered the additional information offered by the social linkage UI as being beneficial to their selection of which network service to use. However, participants tended to evaluate this extra information only at a shallow level. Even in cases where the distribution of users associated with one service was as high as 90%, average user confidence in the genuineness of that service never exceeded 75%.

The second hypothesis of the experiment, that *for a given network, as the number of users reported by the system increases or decreases relative to the number of plausible* physical *users (i.e. that the user can physically see in their immediate environment), confidence in the genuineness of that network will reduce* was *not* supported. Despite participants considering the additional information presented by the H2 conditions (the actual numbers of nearby people that were connected to each of the two services examined) as being more useful than the proportional distribution presented in the H1 conditions, this had no significant effect upon the degree to which they considered either service as being more genuine than the other. Examination of participant reasoning suggested a number of reasons why this was so. A substantial number of participants clearly understood aspects of radio phenomena relating to wireless communication, particularly with respect to its ability to permeate outside their immediate physical surroundings. This understanding was subsequently used to present good reasons why the numbers presented by the system might not entirely tally with the numbers of plausible wireless service users that they could directly observe (e.g. the system was able to detect devices outside of the participants field of view).

In terms of the role that social-linkage based evidence was hoped to play, that participants felt that they could not rely upon the number of plausible wireless service users that they could see in order to validate the numbers as reported by the system was disappointing. This effect was compounded somewhat by the substantial disparity (an average of -15%) between the actual number of plausible users physically present (as calculated by the researcher) and the same number as calculated by the participants. However, that participants considered the less directly verifiable presentation of the proportional distribution of users associated with each network service (i.e. the H1 conditions) as being relevant to their confidence in the genuineness of one service over another is in itself a promising finding. Further refining of the

social linkage protocol would seek to expand on this finding, reducing reliance upon directly verifiable accuracy and instead exploring ways of supporting socially-derived evidence of service usage through other means. Suggestions for future work include the ability for users of a particular service to rate / comment upon that service for the benefit of future users, and / or the ability for the system to derive historical usage patterns associated with a given service (thus providing evidence of how well established that service is within a particular physical location).

6.7 Chapter summary

In this chapter, the role of a third potential linkage channel *social linkage* was examined in terms of its effect upon perceptions of situated service trustworthiness. To facilitate this investigation, a typical wireless network service discovery protocol was extended to provide information about the numbers of users that were associated with each detected service. In doing so, the experiment sought to address one of the two characteristics of pervasive computing usage as identified in the introduction to the thesis: a need for improved methods of *service discovery*.

In a semi-laboratory based experiment, the social-linkage enhanced service discovery protocol was evaluated by a number of participants in a scenario whereby two essentially identical services were made available, with only one known to be actually genuine. The results of the experiment suggested that, while providing information that people found to be useful in aiding their decisions as to which service to choose, the role of social-linkage as a user-verifiable evidential cue was found to be somewhat less effective than the other forms of linkage examined thus far.

This experiment concludes the exploration of the *linkage* concept. In the next chapter, a final experiment is described in which the effectiveness of using *linkage* as a means of increasing the trustworthiness of a situated service was tested in a scenario whereby participants would actually have to invest their trust in a service, as opposed to merely stating their intention to trust a service.

Chapter 7

Measuring trust investment in an experimental setting: *WiFi Phishing*

> "Risk, or meaningful personal investment is a prerequisite of trust. The need for trust only arises in risky situations, and the trustor must be cognizant of the risks involved".
>
> - Morton Deutsch, *The resolution of conflict: constructive and destructive processes*, 1973 [35].

7.1 Chapter overview

In the previous two experiments, situated service use-case scenarios were simulated in a semi-laboratory setting. Each experiment focussed on a particular characteristic of pervasive computing that was considered as requiring some degree of human-computer trust. These characteristics were *secure ad-hoc device association* (chapter four) and *service discovery* (chapter five). In both experiments, various artefacts in the physical world were leveraged to offer evidence that a given service could be considered as being trustworthy. This effect was achieved through using those artefacts to show that a given service was linked in some way to the establishment in which it was found. Thus far however, the experiments conducted have all measured the effectiveness of linkage-based evidential cues through measurements of participant stated *intention* to trust.

As many researchers, including Malhotra (2004) [88] have noted, *trusting intention* (a psychological state) and *trusting acts* (the actual behaviour of investing trust) are two different things. Thus, an apparent / stated willingness to invest trust in a situated service might not necessarily translate to an actual investment of trust. To truly investigate the degree

to which linkage could form a useful means of enabling well-placed user trust in situated services, measurements of actual trust-investment behaviour as a result of their presence would be ideal. However, for a number of reasons that will discussed in this chapter, observing and recording the true investiture of trust in a controlled experimental setting is methodologically challenging.

In this chapter, an experiment was designed that sought to investigate actual trust investment behaviour in a real-life situated service usage scenario. To achieve this, a spoofed situated service (a WiFi 'hotspot') was deployed as an experimental device in a number of real-world cafes, and data collection was undertaken without the experimenter being physically present (i.e. as an 'unattended' study). Participants who engaged with the situated service were not made aware that they were participating in a controlled experiment. In order for the WiFi 'hotspot' service to be used, participants in the experiment were required to engage in a user-verification process whereby they were obliged to supply the service with some personal information (their mobile phone number). As the situated service developed for the experiment involved this potentially risky investment in order that it be used (and as the participant would have no prior experience with the service), users would need to trust the system in order to obtain the benefits that it purported to offer. As an experimental device, the researcher would be able to capture and log instances of user trust-investment behaviour though observing those instances where users chose to complete the verification process.

7.2 Introduction

'Phishing', a relatively novel variation on established methods of confidence trickery, is the practice of attempting to acquire personal or sensitive information fraudulently through the use of electronic communications. 'Phishing' communications, commonly email based, are often engineered to appear trustworthy by closely resembling real communications from trusted people or institutions. 'Phishing' attacks often involve mimicking the on-line communications of the banking/financial services, whereby access credentials are captured and subsequently used for monetary fraud. Such attacks have been shown by experimental research to be worryingly effective (Grazioli & Jarvenpaa, 2000 [57], Dhamija et al, 2006 [38], Downs et al, 2006 [40], Sheng et al, 2007 [134]).

A common component of phishing attacks that are deployed via email is a hyperlink to a corresponding website. The website (like the email itself) will purport to be a familiar bona-fide service, but will in fact be a masquerade. With regard to systems deployed using the WWW, fully functional but completely bogus websites can easily be deployed. In urban spaces, with the continued growth of cellular phone and WiFi service provision (e.g. Wifi 'hotspots') in populated areas, the potential for their fraudulent abuse by 'phishers' is increased for a number of reasons. As the source of a novel wireless service is often neither easily ascertained nor easily verifiable, by using a laptop running as a standalone web-server, the resourceful 'phisher' can easily deploy a fully functional digital service to which people can connect while he himself remains undetected.

7.2.1 Why phishing works

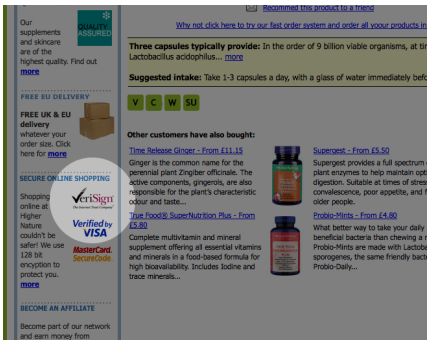
The art of 'phishing' relies on convincing people that a communication is genuine when it is not. As such, the success of any phishing attempt is largely dependent on its victims failing to detect the deception. Within 'phishing' type communications there are likely to be subtle cues as to its deceptive nature. Increasing the effectiveness of methods by which user attention can be drawn to such cues is often considered key to the prevention of such attacks.

That user attention can and often should be more effectively directed is an issue of interest to technology researchers and designers of user interfaces in general. Rutter (2001) [126], discussing technology use in general, found that when people approach a new experience, they often tend to automatically apply rules that have governed similar experiences in a similar domain. This phenomenon is discussed elsewhere in the social sciences, including the *set effect* (A.K.A. *Einstellung*), a tendency to initially and automatically deploy well-tested strategies in situations of ambiguity (Good, 2000 [54]). The *set effect* can be considered as a type of 'cognitive inertia' – an autonomic response designed to short-cut the cognitive effort required in closely evaluating every situation that a person comes to face. The use of a masquerade to perform a phishing attack can be thought of as taking advantage of this effect as it relies upon the user failing to pay close attention to something due to its familiarity.

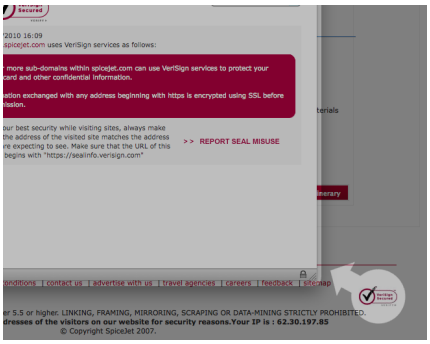
Designers of e-commerce and other web-based service providers that rely on the use of sensitive information have sought to reduce the vulnerability of their systems to phishing attacks in a variety of ways. Many of their efforts have sought to increase the number of cues that are made available, and particularly how those cues can be made salient to the user such that their attention is drawn upon more effectively. However, there remains much inconsistency between the techniques used by the most popular Internet browsers, and the general effectiveness of many of the solutions presented thus far have been found to be less than ideal. Browser-based security indicators provide one such example. Browser-based security indicators are typically made manifest within the 'chrome' of the browser. The 'chrome' refers to the space surrounding the content of the website being viewed, and as such is part of the browser software rather than the website itself. In many modern browsers, information relevant to the security of the website that is currently loaded are displayed prominently, alongside the address and / or status bar of the browser's chrome. However, as Dhamija et al found when examining how and why phishing works (Dhamija, 2006) [38], a significant proportion of people (23%) do not pay close enough attention (if indeed any attention at all) to chrome-based indicators of website security. When assessing a website on initial encounter, 23% of Dhamjia et al's participants determined the legitimacy by the content of the site only. 36% did so by a combination of the website content and its domain name. The better engineered spoof websites that were designed for the study successfully fooled 90% of the people that were tested. Alternative efforts to increase the salience of chrome based security mechanisms have been reviewed in Dhamija et al (2005) [37].

Within the website content itself, graphically-based *trust seals* are commonly used to offer a visual cue that that website has been checked and validated by an independent institution. Two examples are *Verisign* and *Trust-e*). By following a hyperlink that is embedded

within the *trust seal*, users of the website can further validate this information by checking corresponding information on the independent institution's website. However, a casual examination of a number of websites that contain such a seal found only a few examples of where this additional verification function was actually present (figure 7.1).



(a) *Verisign* trust seal without corresponding link to supporting evidence



(b) *Verisign* trust seal with corresponding link to Verisign (link target shown as overlying window)

Figure 7.1: Two examples of the use of 'trust seal' graphical cues in e-commerce websites. In the first example, the seal exists only as a graphic, with no immediate means for the user to verify it via Verisign.

7.3 The design of an experiment to deliberately 'phish' situated service users

In this experiment, a bogus Wifi 'hotspot' service was installed in a number of real-world cafes that were open to the public. Though the management of the venues used would be aware of the existence of the service, no information about the service would be offered to customers. The service would be free to discover and use.

As has been observed in the experiments described thus far, whether a person considers a situated service as being worthy of their trust involves a complex mix of both cognitive and affective reasoning. Informing this reasoning are evidential cues, of which the offering of links (*linkage*) between the physical and digital world appears to be one such cue. Thus far, through evaluating three forms of linkage (*physical*, *virtual* and *social*), a number of mechanisms have been identified that appear to increase the trustworthiness of a situated service, with a corresponding increase in stated *intentions* to trust.

Trust is a complex and multi-faceted concept that poses a significant challenge to researchers keen to isolate the phenomenon in a controlled experimental setting. A substantial contribution to this problem is due to the fact that, as researchers including Deriaz (2006) [33] have noted, notions of trust and risk are dissociable. To invoke a sense of risk in a controlled environment, experimental trust research, particularly in economics have generally relied upon the use of laboratory based *trust games*. Typically, trust games are variances of the *prisoner's dilemma* protocol, of which one example is the *investment game* (see section

2.4.3). To further engender risk, several trust-game research programmes have used real money and thus created the potential for true participant loss. However, as trust game experiments are typically laboratory based, they are inherently susceptible to contamination by the *white coat effect*, i.e. the (unintentional) authority ascribed to the experimenter by the participant. In the experiment presented in this chapter, possible *white coat* effects were removed by conducting the experiment 'unattended' - i.e. without the experimenter being physically present. In the experiment, risk would be simulated by requiring participants to volunteer personal information in order to use a situated service. Ideally, the form of that personal information would be highly sensitive, such as a participant's credit card details. However, for reasons of security and the potential for participant discomfort, pursuit of this option was abandoned. Instead, the experiment would require the entering of a personal mobile phone number. A personal mobile phone number was thought to offer a suitable alternative example of personal information that still carried a real potential risk of abuse.

7.3.1 Research questions and hypotheses

The practicalities of installing an essentially rogue wireless situated service in a real-world working cafe meant that, unlike in the previous experiments, the potential use of additional physical artefacts (such as wall posters, leaflets and screens) was severely constrained. Instead, a physical linkage cue was created though means of an image embedded in the website that was presented to users as they attempted to connect to the service.

The primary prediction of the experiment was that the presence (within the website) of a highly salient photograph that uniquely represented the user's current location would increase the likelihood of the user 'trusting' the website enough to supply some personal information (in this case their personal mobile phone number). This is as opposed to the same website displaying a photograph of a more generic location that is not identifiable as being unique to the user's current location. By including a location relevant photograph as a salient evidential cue of communication source, it was hypothesized that user uncertainty about the source of the communication would be reduced through offering a link between the service and the venue where it was deployed.

The role of the embedded image was thus to serve as a form of physical linkage, and it would achieve this through means of providing a *locative* cue. A *locative image* is thus described as an image or other form of media that is embedded within the service's content, and which represents the location where the (source of the) service is situated. The presence of a locative image formed hypothesis one:

H1 *Locative hypothesis*: The presence of a locative image cue would increase trust in a Wi-Fi hotspot compared to the presence of an non-locative image.

However, it is also important to examine what is meant by the absence of a locative cue. The idea of an image *not* representing a given location can be broken down into one of the following mutually exclusive categories: **Anti-locative:** an image representing a place that is specifically *unlike* the given location and **a-locative:** an image representing a location that could be in any of many places, including the given location. The presence of an

a-locative cue that, while not inconsistent with the user's location, does not specifically represent that location was considered as having no effect on trust investment. However, the potential effect of an anti-locative image formed the second hypothesis:

H2 *Anti-locative hypothesis*: The presence of an anti-locative image would decrease trust in a Wi-Fi hotspot compared to the presence of an a-locative, or locative counterpart.

7.3.2 Experiment environment and materials

For the purposes of the experiment, a situated service called *Fastnet* was created by the researcher. *Fastnet* purported to be a free wireless internet gateway service (a WiFi 'hotspot'), and was deployed in real-world cafes in two major U.K cities (Bristol and central London). The venues chosen were selected based on their having existing WiFi Internet access points and established WiFi user bases.

7.4 Experimental design

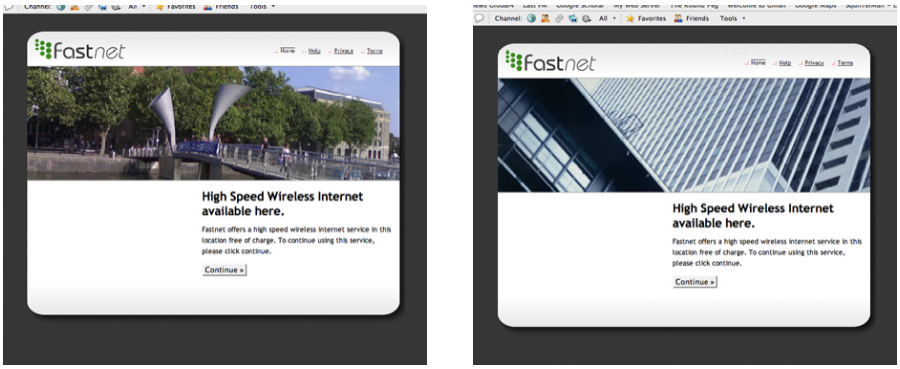
The experiment was a between-subjects design that was conducted without the researcher present during data collection. The requirement that participants be subjected to an intentional deception without their prior consent was considered carefully in terms of the ethical impact this activity might have on participants. To ensure confidentiality and security of user supplied information, the experiment was conducted within SSL secure web space. Participant mobile phone numbers were encrypted at source, transmitted securely and hashed upon arrival at the *Fastnet* web-server(s). Thus, while records of participant mobile phone numbers involved in the experiment were stored (for the purpose of preventing instances of the same participant performing the experiment more than once) they could not be read by the researchers, nor could they be used to identify participants directly. All participants who supplied their mobile phone numbers to the *Fastnet* system were debriefed by the system and multiple modes of contacting the research team were made available. Ethical approval for the study was applied for and granted by the University of Bath department of Psychology.

7.4.1 The independent variable and the creation of experimental materials

Within the experiment, two independent variables (IVs) were manipulated. These were *location* with two levels (Bristol, London) and *image*, also with two levels (local, generic). The levels of *image* are described thus:

Local image: The *Fastnet* service website contained a masthead photograph representing the local area immediately outside the venue used in Bristol (figure 7.2a).

Generic image : The same website, but with the masthead photograph replaced with a photograph representing a generic urban environment (figure 7.2b).



(a) *Fastnet* splash screen displaying the *local* image condition.

(b) *Fastnet* splash screen displaying the *generic* image condition.

Figure 7.2: *Fastnet* splash screen in the *local* and *generic* conditions.

The selection of photographs for the independent variables *image* and *location*

The selection of the photographs used to create the conditions for the IV's *image* and *location* was achieved through a simple photograph ranking exercise that was undertaken with members of the public (at both Bristol and London venues). Participants (Bristol: n = 21[M=12, F=9]) , London: n = 20[M=11, F=9) were asked to rank seven photographs (see appendix D, section D.1) in terms of the statement *this photograph most clearly represents where I am now*. Three of the photographs were of the immediate area adjacent the cafe venue in Bristol (thus *locative* for Bristol, but *anti-locative* for London), three were non-specific photos of urban-style architecture (all *a-locative*), and the remaining image was a 'wildcard' photo that was not directly representative of either location of (a New York street scene that should be considered *anti-locative* in both London and Bristol). The overall highest ranked image in Bristol was used as the *local* image condition and the second lowest ranked image (disregarding the wildcard image) formed the *generic* image condition. Validating the choices made, results from the London participants indicated that the *local* image was seen to be second least representative of that locale, while the *generic* image was found as second most representative.

Both image conditions were tested in both the Bristol and London cafes. However, as the location specific photograph was relevant only to participants completing the experiment in Bristol, to support the hypothesis, it would be expected to find a significant difference in the levels of trust investment across the two image conditions in Bristol only. As neither photograph provided any relevant local cue to participants in London, no difference in the levels of trust invested were expected as result of the presence of one photograph over another. With the exception of the masthead image, the appearance and functionality of the *Fastnet* website was identical at both locations.

Design issues related to unattended experimentation

Participant recruitment: The engendering of risk in the experiment was deliberately heightened by conducting the experiment 'unattended'. As the researcher was not present however, the recruitment of participants was entirely dependent upon visitors to the cafe(s) both discovering and connecting to the *Fastnet* service of their own volition. This posed substantial technical problems in the design of the experiment, not least the problem that many mobile devices would remember services to which they had previously connected and would automatically reconnect to them if they are detected again. As both the cafe venues used had existing WiFi services available, participant engagement in the experiment was limited by their mobile devices failing to offer the option to connect to a different service by default.

Consequently, the reality was that the collection of substantive participant data would take a matter of months rather than days / weeks. This issue of timescale, asides from demanding a degree of patience, had an additional knock-on effect upon the technical aspects of the experiment; the *Fastnet* system had to be robust enough to withstand a significant time period of sustained activation. In order to provide continual service over an extended time frame of several months, and where power sources were potentially variable, the web-servers used to deploy *Fastnet* were based on small footprint Linux-based laptops. This provided a increased degree of stability, and meant that the system could be restored swiftly by non-technical persons present within the cafe(s) when required. Further, to provide the researcher with a commentary on progress, both web-servers also transmitted a report of their status via a 'heartbeat' that was delivered via the cellular network short messaging service (SMS) on a daily basis. As well as providing an early warning of system problems, this also supplied a useful daily report of participant engagement, and the number of participants who had successfully been 'phished'.

Assigning participants to the two levels of IV1 *image*: In terms of system design and implementation, error capture, recovery and access control were all critical factors. To ensure control of access and balanced participant condition assignment, the *Fastnet* system both identified and recorded the unique MAC address of each user's device. This information was then used to assign each user to one of the two levels of IV1 and ensured that multiple access to the site by the same machine resulted in exposure only to the originally assigned condition.

***Fastnet* service website: Aesthetic and functional considerations:** The unattended nature of the experiment demanded a system that could cope with a myriad of potential devices and operating systems that could conceivably attempt to make connection. Interaction with the *Fastnet* service was conducted through a website that presented itself to participants upon connection to the service.

A professional web designer was employed to create the *Fastnet* website. The site itself was deliberately designed to be minimalist in tone, content and color, while adhering to design conventions of sites of a similar kind. To ensure maximum salience, around 50% of the active space of the website was devoted entirely to the experimental manipulation (a photograph), and other imagery usage was minimal. On a functional level, the website code was built

to comply with W3C xHTML web standards, and was subject to an exhaustive testing schedule with a broad range of Internet browsing software that was available at the time. These included Safari, Internet Exporer, Firefox (Macintosh), Internet Explorer, Firefox, Opera (PC) and the proprietary browsers of several mobile devices (mobile phones / Pocket PC based devices). Finally, and in response to the Dhamjia et al (2006[38]) finding that 36% of participants tested on a variety of spoofed 'phishing; websites utilised the domain name as well as the content of the site when making judgements of site legitimacy, our service ran with the domain name 'http://www.fast-net.org'.

7.4.2 Dependent variables / experimental measures

Within the experiment, the dependent variable was 'phishing success', measured in terms of a simple boolean outcome: 'phished' or 'not-phished'. Whether a participant was considered as having been 'phished' was dependent upon them successfully traversing the *Fastnet* service authentication process completely ('phished'), or leaving the process prematurely ('not-phished'). To manage this process, the *Fastnet* service website consisted of eight individual web-pages (see figure 7.1). Four pages formed a 'login' process, three offered help and information (as to the use of the service), and one was used as a 'blocking' page that prevented abuse / multiple logins from the same device. The path through the login process was forced: home - login - password - debrief. Each step of the process is described later in this section. Attempts to subvert the path through the website / jump steps (e.g. through direct URL entry) were automatically redirected to the 'home' page. Links to the help and information pages were available from every page.

Web page	**Function of web page**
Home	Presented the 'Fastnet' service, invited user to login.
Login	Presented the requirement of a valid mobile phone number to. Requires entry of phone number to proceed.
Password	Presented the requirement of the code that the system had sent via SMS to the supplied phone number. Required entry of a valid code to proceed.
Thankyou	Experiment completed with participant 'phished'. Presented debrief.
Blocked	Displayed upon attempt to use a previously used mobile phone number.
Terms	Displayed fictitious terms and conditions information, based upon actual websites offering the same service.
Privacy	Diplayed a fictitious privacy policy, based upon actual websites offering the same service .
Help	Displayed basic website usage information.

Table 7.1: Summary of the *Fastnet* website structure.

Step 1: 'Splash screen' - *welcome to Fastnet*

Upon connection to the *Fastnet* service, the MAC address of the connecting device was recorded and used to assign the participant to one of the two image conditions (see figure 7.2).

In the event of a repeat visit, assuming the same device was being used, any given participant would only ever be presented with their originally assigned image. When an Internet browser was opened, participants were automatically presented with an introductory splash screen that introduced *Fastnet* as a free wireless internet gateway with details of how it could be accessed. Any attempt by the user to bypass or subvert the *Fastnet* website (e.g. through direct entry of a URL) resulted in the user being redirected back to the splash screen.

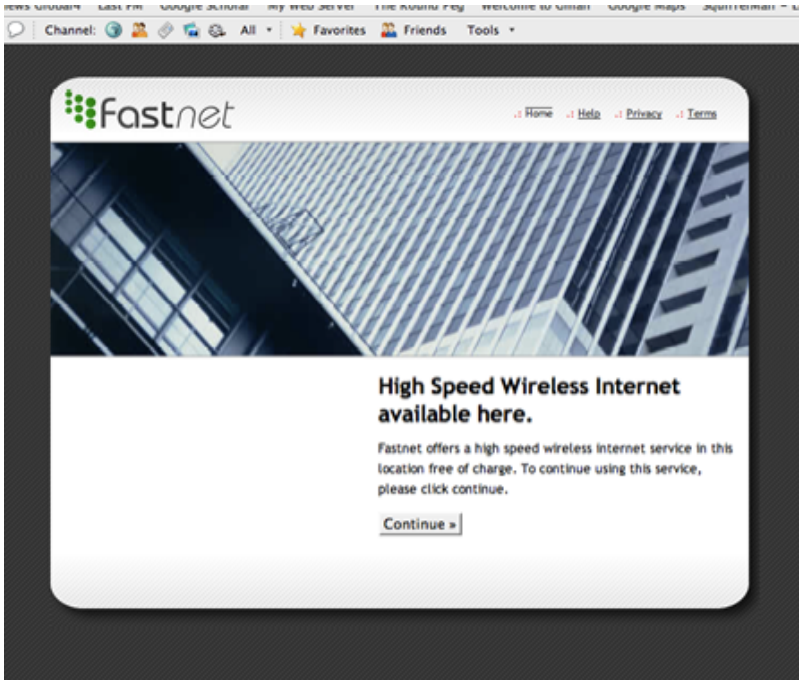


Figure 7.3: *Fastnet* access: Step 1: *Splash screen.*

Step 2: Login to use *Fastnet - please supply your mobile phone number*

Participants who chose to continue were then asked for their mobile phone number in order that they authenticate themselves before being able to use the service (figure 7.4). An explanation was offered that the service offered fast access in return for a degree of accountability on the part of its users.

Step 3: Authentication - *please enter your unique passkey*

Upon submission of a valid mobile phone number, participants were informed that they would shortly receive a unique personal identification (PIN) number via their mobile phone (SMS *Short Messaging Service* text messaging was used for this process). Participants were told that they would need to enter this PIN into the *Fastnet* website in order to complete their authentication and start using the service (figure 7.5) .

Figure 7.4: *Fastnet* access: Step 2: *Login*.

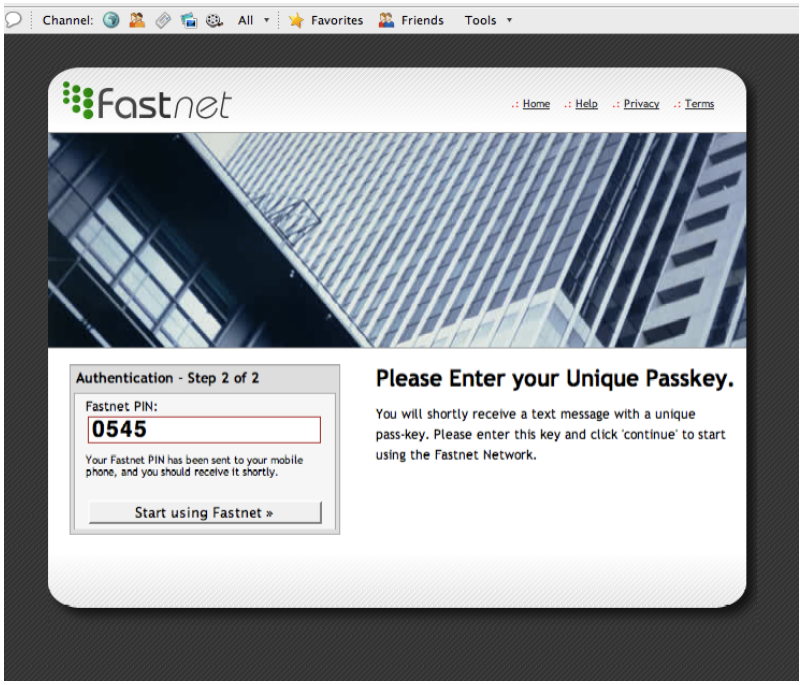


Figure 7.5: *Fastnet* access: Step 3: *Authenticate*.

Step 4: Participant debriefing

At the point at which the correct PIN had been entered and submitted, the experiment ended and the participant was debriefed by the website (figure 7.6). The debrief including details

of the experiment, the security of participant information supplied during the deception and the contact details of the experimenter and project staff.

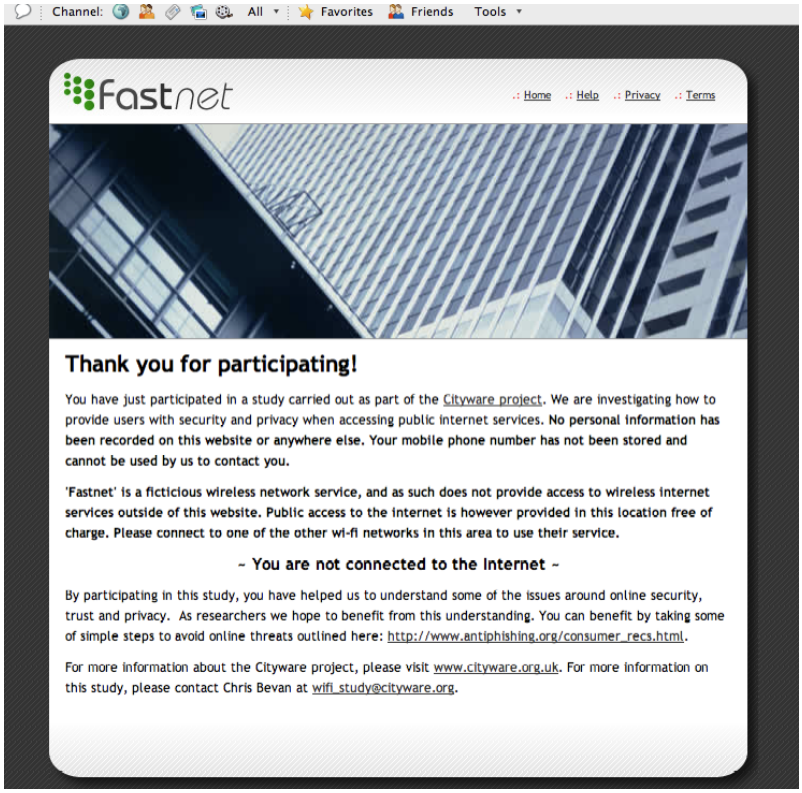


Figure 7.6: *Fastnet* access: Step 4: *Debriefing*.

7.4.3 Participants

The participants were 361 members of the public (n[Bristol] = 247, n[London] = 114), identifiable only by the unique MAC address of the devices they had used to connect to the *Fastnet* system. Age and gender distributions were unknown.

7.5 Results

The results presented in this section were generated through a combination of logs generated by the *Fastnet* web-server(s) and additional logs generated by the *Fastnet* system itself. Data was collected over a period of 29 weeks between October 2006 and July 2007.

7.5.1 Patterns of site access and instances of 'phishing'

The distributions of unique connections (as provided by unique MAC addresses) and phishing events recorded by *Fastnet* for each of our two locations are presented in figures 7.7 and 7.8 respectively.

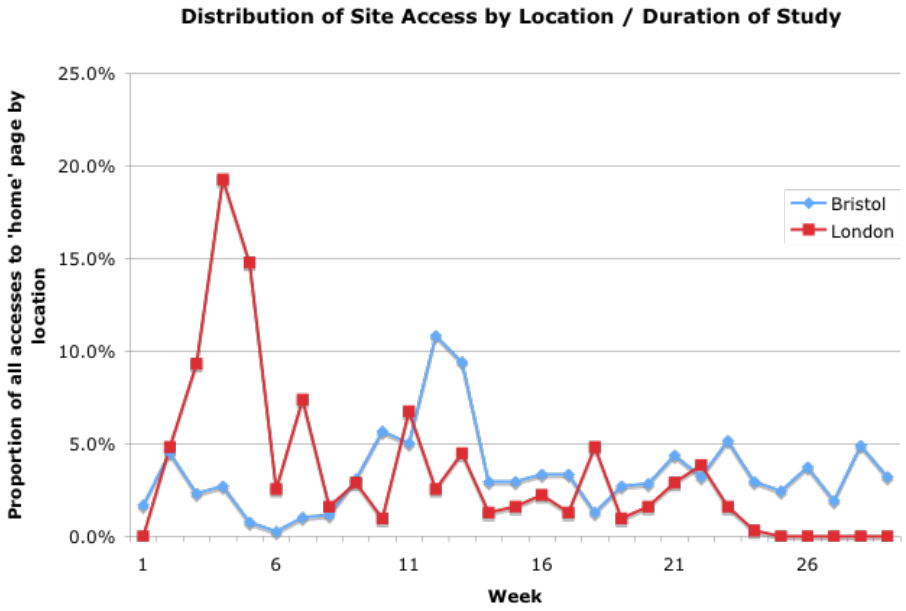


Figure 7.7: Distribution of connections made to the *fastnet* website recorded over the duration of the study.

The general phishing success rate (independent of location) was approximately 32%. Around half (53%) of the participants who failed to be 'phished' exited the site without progression past the splash screen. 29% exited at *login*, 12% at *password* and the remaining 6% left from one of the three *help and information* pages. 80% of participants made only one visit to *Fastnet*. The spike in phishing rates noted at the Bristol venue between weeks four and five coincided with the period leading up to Christmas 2006, where it was assumed that the venue (being adjacent to the city shopping district) had encountered much higher than usual customer traffic due to the Christmas season.

7.5.2 The effect of *location* and *image*

Any participant receiving the debriefing page was considered as indicating a successful 'phishing' (the system otherwise prevented access to that webpage). Raw counts of accesses made to the *debriefing* page were calculated for each participant by location / image condition and are presented in table 7.2. As any given participant could only be 'phished' once, any repeat hits (e.g. as might be incurred by a participant deliberately or accidentally

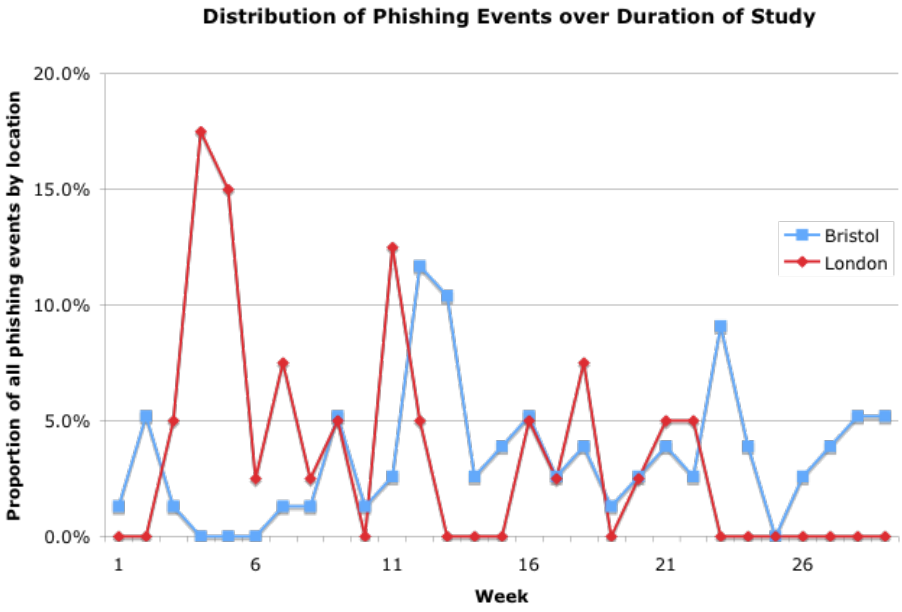


Figure 7.8: Distribution of phishing events recorded over the duration of the study.

reloading the web page) by the same participants on the debrief page were removed from subsequent analysis.

Location	Image	Total participants	Phished	Not phished
Bristol	Generic	122	36 (29.5%)	86 (70.5%)
	Local	125	41 (32.8%)	84 (67.2%)
	Total	**247**	**77 (31.2%)**	**170 (68.8%)**
London	Generic	59	26 (41.1%)	33 (55.9%)
	Local	55	13 (23.6%)	42 (76.4%)
	Total	**114**	**39 (34.2%)**	**75 (65.8%)**
Grand total		**361**	**116 (32.1%)**	**245 (67.9%)**

Table 7.2: 'Phishing' success rates by *location* and *image*.

Comparative 'phishing' rates across the location and image conditions are presented as a plot in figure 7.9. A three-way loglinear analysis produced a final model that retained all effects. Results showed no significant main effect of either *location* ($x^2(1)=0.407$, $p=0.523$ n.s) or of *image* ($x^2(1)=2.639$, $p=0.104$ n.s). However, there was a significant interaction between *location* and *image* ($x^2(1)=4.886$, $p=0.027$), whereby phishing rates were lower for *image LBristol* and higher for *image NLBristol* in the London location only.

To further investigate the interaction between *location* and *image* and their effect upon 'phishing' success, additional chi-square tests were performed on the *image* and *phished* variables separately for both locations. For London, there was a significant association

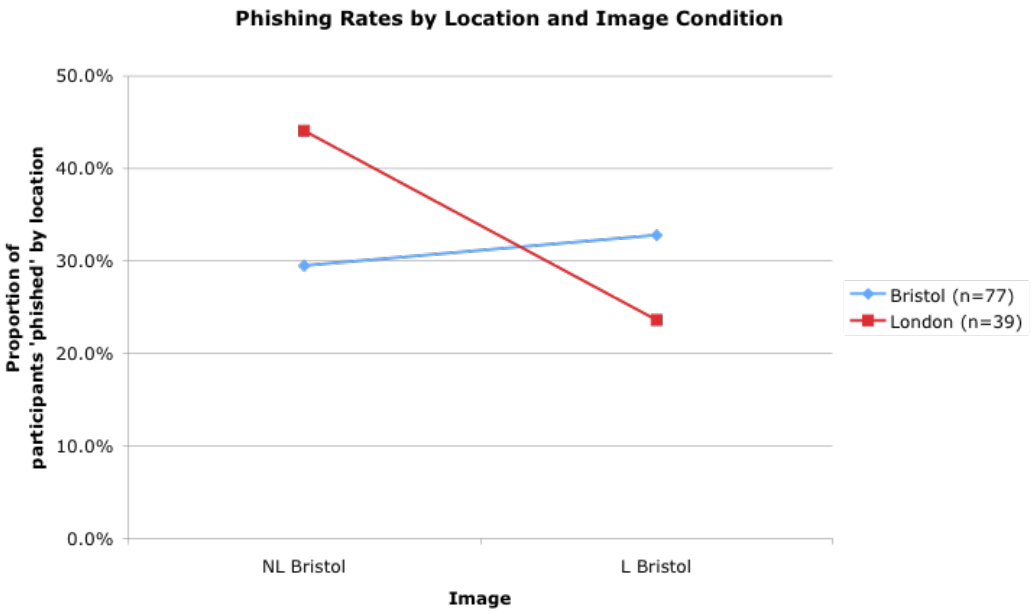


Figure 7.9: Proportional 'phishing' rates by *location* and *image*.

between the image shown and the likelihood of subsequent 'phishing', $x^2(1)=4.41$, $p=0.036$ (using Yates' continuity correction); this was not true of Bristol, $x^2(1)=0.177$, $p=0.674$ n.s (using Yates' continuity correction). Odds ratios indicated that participants in London were 2.55 times more likely to be 'phished' when presented with the image *NLBristol* than if presented with image *LBristol*.

The analysis revealed a difference in the incidence of 'phishing' for Bristol and London: while participants in Bristol were equally susceptible to being 'phished' regardless of the image displayed, participants in London were much less susceptible to being 'phished' when presented with image *LBristol* (the image selected as a locative cue in Bristol) than with *NLBristol* (the image selected as being a-locative in both locations).

7.6 Discussion and conclusions

The primary intent of the experiment presented in this chapter was to produce a means by which the measurement of trust was as expressed through actual trust-investment behaviours rather than merely through the asserted intention to trust.

The results of the WiFi phishing experiment gave rise to two general findings. These were: 1) a substantial rate of successful 'phishings' and 2) evidence for location as being relevant for trustworthiness in a situated service scenario. Both findings are discussed in turn.

A substantial rate of successful phishings. In both locations examined, irrespective of the image condition, around a third of people who encountered the *Fastnet* service trusted it with their mobile phone number. We might assume therefore that similar numbers of people might have trusted that service even if nothing was required from them in order to use it. The fact that so many users entered their mobile phone number into *Fastnet* suggests that Wifi providers should consider protecting their users. Though an attacker might conceivably utilise the users mobile phone number for malicious ends, the active connection to *Fastnet* could also provide sufficient means for the attacker to eavesdrop all communications made whilst using the service and perhaps install malware on the users machine. The results of this experiment therefore has implications for the design of situated services such as WiFi hotspots: Consumers need to be protected from mistakenly trusting spoofed services, and designers need to consider and avoid the potential for distrust as a barrier to use of legitimate services.

Evidence for location as being relevant for trustworthiness in a situated service scenario. A significant result was observed whereby the inclusion of an image that was considered anti-locative led to significantly less trust (and by extension less successful 'phishings') than an *a-locative* or *locative* image. The interaction is presented in figure 7.10.

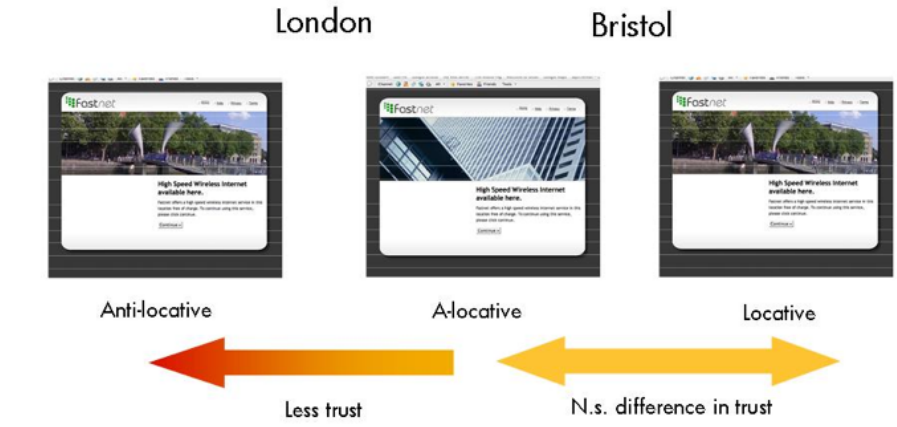


Figure 7.10: Support for the a-locative hypothesis.

While the experiment was found to support the *anti-locative* hypothesis (**H2**), it failed to support the *locative* hypothesis (**H1**). Unfortunately, the unattended nature of the experiment meant that developing a deeper understanding of these results was rendered impossible. As the researcher could not engage with the participants, a more detailed analysis of participant reasoning that might have shed substantial light upon the value of location as a cue to trustworthiness was not available. Thus, while it was considered that the differences in 'phishing' rates that supported **H2** were likely due to the anti-locative properties of the images shown, the possibility that some other attribute created this effect cannot be easily dismissed. However, it was considered that, as the evidence that could conceivably be used in what was a real-life working cafe was limited, the result was at least compelling enough to warrant further investigation.

The use of unattended experimentation in this instance had serious limitations and the re-

searcher had to accept the consequence that interpretation of participant reasoning with no experimenter present in-situ would be limited to inference alone. The use of an unattended field study was chosen specifically for its potential to create a scenario whereby participants would be exposed to (at least as far as they know) real risk. As the presence of risk is necessary in any scenario where trust investment can conceivably occur, it was felt that this was critical to the experimental design. Further, the absence of the experimenter served to strengthen this illusion while also reducing the possibility of the unintentional influence upon participant reasoning and decision making. Indeed, it is difficult to propose a way in which this compromise might have been avoided. As the experiment utilised deception, obtaining participant feedback via the *Fastnet* system was difficult for a number of reasons; participants who failed to be 'phished' could not be alerted to the true nature of the experiment, and participants who were 'phished' could well be assumed as being not particularly disposed to help the researcher. Further, as the WiFi hotspot service provided by *Fastnet* was in actuality a ruse (no actual connection to the wider Internet could possibly be reached), although provision of means by which participants who did get 'phished' could contact the researcher, no communications were ever received.

7.7 Chapter summary

In this chapter, addressing a commonly cited shortcoming of empirical trust research, an experiment was designed to measure actual trust-investment behaviour in a real-life situated service (a WiFi hotspot). Within the experiment, a physical-linkage cue was embedded within a website that formed the service connection protocol. The physical-linkage cue in this case was an image that related to the physical surroundings of the user to a variety of degrees.

The results of the experiment gave rise to some interesting findings. The first of these findings was a surprisingly high degree of initial-situational trust with a service that demanded personal information up-front and without assurances. In order that they could obtain the benefits it purported to provide (free access to the Internet), a third of people who (unknowingly) participated in the experiment apparently trusted the service enough to volunteer their personal mobile phone number to it.

Further, an effect of linkage based evidence as a means of imbuing a sense of situated service trustworthiness was also found to occur, albeit to a degree that was weaker than was found in the experiments previously conducted. In situations where participants were exposed to an image that clearly did not represent their current physical surroundings, trust in that service as a whole was significantly reduced, relative to an image that either directly or indirectly represented their current location.

Chapter 8

Conclusions and further work

8.1 Introduction

Consequent of the less visible, wire-free nature of pervasive technologies such as *situated services*, the number of means by which such services can convey their trustworthiness to their potential users is reduced. This thesis sought to address the processes and mechanisms by which people assess the trustworthiness of a pervasive situated service upon their first encounter. Using a mixed methods research plan in both laboratory and field, this process developed an understanding of how people respond to the potential risks and threats that are associated with this type of system. Through this understanding, the research then continued by both developing and evaluating methods by which users of situated services could be better informed about the trustworthiness of the services they encountered. In conducting this research, the main aim was to help inform pervasive systems designers as to how they might best incorporate those processes and mechanisms within their future designs.

The principle prediction to which the thesis was based was that people, in the absence of prior knowledge, would seek to identify and use contextual evidence that was available in their immediate environment as a means of assessing the potential trustworthiness of a new situated service. This prediction was derived from Mcknight et al's (1998) [101] model of initial-situational trust, in which the *identification and assessment of available evidence as to the intentions and competence of the trustee* is presented as an antecedent factor in the generation of initial-situational trust investment behaviour. Over three experiments, a conceptual solution to this problem was posited and tested: '*linkage*'. '*Linkage*' referred to means by which user perceptions of the trustworthiness of a given situated service could be increased by providing human perceptible links between the digital service and the immediate environment of the user (e.g. Fogg (2001 [46], Riegelsberger & Sasse (2001) [117], Giddens (1990) [51]). Through providing user-verifiable evidence of this link, the thesis predicted that user perceptions as to the trustworthiness of a given wireless digital service would be increased.

In the experimental component of the thesis, three distinct channels of linkage were identified, explored and tested for their effectiveness in increasing initial-situational trust in a novel situated service (a WiFi 'hotspot'). The three channels were: *physical* linkage, *virtual* linkage and *social* linkage. *Physical* linkage referred to the use of artefacts that were connected in a simple physical sense to the environment, in such a way that it would be difficult for an attacker to contrive (e.g. mounted to a wall). *Virtual* linkage referred to the use of interactions with a point of trust that an attacker could not easily contrive. Finally, *social* linkage referred to the use of physical persons present as a means of offering evidence that a situated service was under active use. The results of these experiments have several implications for the design and deployment of pervasive situated services in urban areas. This chapter will continue firstly by providing an overview of the experimental results gathered. The contributions of the thesis are then presented and, in conclusion, the research conducted so far will be reflected upon, with potential avenues for further work discussed.

8.2 Overview of the experimental results

8.2.1 Current user experience and behaviour with technologies that involve issues of personal privacy and security

In chapter three, initial investigations into user experience and behaviour relating to their current risk-relevant ICT service usage were conducted through the development and deployment of an online questionnaire. From the results of the questionnaire, respondents were found as having a high level of general experience using technology to conduct activities that involved risks to their personal / private data. Respondents presented themselves as being highly security conscious and protective of their personal / private data, many having had personal experience of Internet-related malicious behaviour such as phishing, credit card fraud and virus infection. While respondents on the whole appeared content to engage in behaviours that would make their personal / private data vulnerable, they would only do so where a clear benefit could be seen to using a particular technology. Technologies that involved invasion of user security / privacy were carefully weighed up in terms of how they were perceived in terms of their relative cost / benefit.

In terms of their experience with technologies that could be considered as possessing characteristics of pervasive computing, despite much experience of using home-based networks and personal computing equipment, respondents had far less experience of conducting risk-relevant technology-based activities in public spaces. Respondents were also found as having less experience still with conducting such activities using non-personal or shared equipment. Thus it was assumed that, with a security conscious mindset but little in the way of past experience, if exposed to a scenario that might be considered typical for pervasive computing (e.g. multiple co-present services), for users to trust those services fully, they would require additional reassurances that the services they chose to use were both genuine and secure.

8.2.2 The effect of physical/digital world linkage evidence as a means of increasing user perceptions of situated service trustworthiness

In chapter five, a scenario was created whereby several co-present situated services were made simultaneously available in a real-world cafe environment, thus more closely reflecting how real-world pervasive computing services might be encountered. Using this scenario, a semi-laboratory based experiment was conducted to examine a specific characteristic and problem of pervasive computing usage: the need for *secure ad-hoc device association* (e.g. Kindberg (2003) [76]).

Participants were invited to assess six independent situated services (WiFi 'hotspots' offering free wireless Internet access) in terms of the degree to which they were confident that each service was genuinely offered by the 'management' of the cafe. To assist them in this process, each service was assigned its own method of authenticating itself as being genuinely offered by the cafe management. Of the six services evaluated, one offered no form of authentication (a control condition), three offered a simple password-based authentication and the remainder used more sophisticated interactive authentication processes. An appreciation of the degree to which participants considered each service as being trustworthy was collected through three measures (obtained both quantitatively and qualitatively): '*genuineness*', '*trustworthiness*' and '*security*'.

In the three password-based conditions, the password required to authenticate the system as being genuine was made available to the participant using artefacts that were physically linked (i.e. in terms of their attachment to the cafe in a simple physical sense) to the cafe to various degrees. These artefacts were, in ascending degrees of physical-linkage, a leaflet, a wall mounted poster and as displayed on a wall-mounted LCD screen. Measures of user confidence in the genuineness, trustworthiness and security of an associated situated service were all found to increase with the degree to which participants considered the source of evidence that was associated with that service as being physically fixed. Further, examination of the qualitative results indicated that several factors were important to participant reasoning about physical fixedness. These were *physical attachment*, the degree to which an artefact was physically attached to some part of the cafe, *legitimacy*, the degree to which participants considered an artefact to be provided by the cafe, *visibility* the degree to which an artefact was in public view (and thus potentially less susceptible to an attacker being able to subvert that artefact undetected) and *accessibility*, the degree to which participants considered the position of the artefact as being accessible by patrons. However, while perceptions of the relative fixedness of the leaflet and wall-mounted poster conditions were sufficiently different to invoke a significant difference in participant confidence, the wall-mounted poster and LCD display-based password conditions were not considered as being any more or less physical linked to the cafe, at least as according to the aspects of physical-linkage that the participants considered.

In the remaining two conditions, two more novel network connection protocols *Synchronisation* and *Interlock* were developed that each allowed the user to verify the authenticity of their associated service through use of a virtually-linked interaction between their laptop and a wall-mounted LCD screen. In the *Synchronisation* protocol, an interaction was

created between the screen of the user's device and the public LCD screen, in which both screens displayed a synchronised series of protocol-generated images. By comparing the degree to which the images on both sets agreed (both in terms of their appearance and the degree to which both screens displayed the images in sync with one another), users were offered visually-verifiable evidence that the two devices are communicating with one another directly. In the *Interlock* protocol, a user-verifiable out-of-band communication channel was used to allow users to generate, communicate and finally compare a number of items of information between their device and the public screen. A sophisticated cryptographic technique underpinned the interaction that could protect the user from communication eavesdropping providing the interaction was performed correctly.

The degree to which each of these protocols was afforded a value of virtual-linkage was dependent upon the degree to which each protocol was able to offer the user protection against two types of communication threat: *evil-twin* and *man-in-the-middle*. As only the *Interlock* protocol could feasibly protect against both, it was considered to be of higher virtual-linkage value than *Synchronisation*. These virtually-linked protocols, when evaluated against the password-based authentication methods, were found to be highly effective at providing user confidence, affording increases in perceptions of trustworthiness, genuineness and security that were significantly above those observed using using physically-fixed artefacts alone. Examination of the qualitative results further indicated that several factors were important to participant reasoning related to virtual-fixedness. These factors were: *Complexity*, the amount of resources that were perceived as being required to subvert the interaction, *synchronicity*, the agreement in time and content between the display and the laptop and *causality*, the degree of cause and effect within the interactivity. However, despite differences in the actual security value of the *Synchronisation* and *Interlock* protocols, users, finding it hard to compare these two unfamiliar and relatively complex protocols, considered both protocols to be equally valuable across all of the measures that were tested.

8.2.3 Socially-derived evidence as a means of increasing user perceptions of situated service trustworthiness

In chapter six, A scenario was created to examine the effect of a second characteristic of pervasive computing service usage: the need to actually locate wireless services (*service discovery*, e.g. Zhu et al, 2005 [158]). Current wireless network service discovery protocols, most notably those found on the most common desktop operating systems, were considered as lacking sufficient information as to the potential trustworthiness of the services that they detected.

In response to this problem, the role that other persons present in a space who might conceivably be using a local wireless networking service was considered. This lead to the positing and empirical investigation of a third potential linkage channel termed *social-linkage*. To examine the effect of a social-linkage channel in a scenario that would require initial-situational trust, a semi-laboratory based experiment was conducted in which two wireless 'hotspot' services, each with a similar name, were presented to participants to choose from (with the understanding that only one of the services was definitely genuine). To assist them in their choice, an existing wireless network discovery protocol was extended to include additional

socially-linked information: specifically the current number of users that were associated with each local wireless networking service detected.

The provision of a system derived representation of user numbers as currently associated with each detectable service was well received by participants. Most participants found the inclusion of socially-derived information at the point of service discovery useful in their decision making about the relative genuineness of one service over another. In the two-service scenario examined in the experiment, as the proportion of users assigned to one service rose above 60%, there was a corresponding significant increase in reported perceptions of 'genuineness' for that service. Further, reported perceptions of service 'genuineness' continued to rise significantly in with each 10% increase in proportion distribution above 60%. Below 60%, however, rejection of both services appeared to be the most favoured outcome.

Despite these findings however, while participants in the experiment considered socially-linked information useful, the results of the experiment failed to fully support the notion of social-linkage. Very few participants appeared to make any direct comparison between the numbers of plausible service users (e.g. laptop users) who were physically present and the numbers as presented to them by their laptop. When directed to do so, participants tended to underestimate the true number of laptop users by around 15%. However, participant reasoning as to why they failed to take this evidence into account was found to be generally sound, if occasionally inaccurate in terms of the technologies involved. When numbers reported by the social-linkage system were considered as being low compared to the number of plausible laptop users physically present, participants felt that this could be explained by connections to a wireless network service being sustainable only through the use of Internet browsing software (and thus people not actively using their Internet browsers were not connected to any wireless networking service). Conversely, when numbers reported by the system were high (compared to the number of plausible laptop users physically present), participants would rationalise this by considering the nature of radio connections, coming to the perfectly feasible conclusion that the system was detecting people who, while proximate to them, were not directly visible (e.g. upstairs or outside the room in which they were sitting).

8.2.4 The effect of linkage-based evidence upon actual trust investment behaviour

A commonly cited shortcoming of much empirical trust research is that, while the collection of measures of stated *intention to trust* (as was performed in the two experiments previously described) can provide a useful means of examining specific trusting attitudes, those attitudes do not necessarily translate into an eventual *trust-investment* behaviour (e.g. Malhotra (2004) [88]). Thus, concluding the experimental component of the thesis, in chapter seven, a field-based experimental methodology was developed and deployed in which actual initial-situational trust-investment behaviour with a situated service could arguably be measured in an experimental setting. The experiment in chapter six had two roles. First, it had to provide a scenario within which a trust-investment behaviour could feasibly occur, and to capture that behaviour should it indeed occur. Second, the experiment would need to measure the effect that some form of linkage-based evidence might have upon any

trust-investment behaviour that was found to have occurred.

To facilitate actual trust-investment behaviour in a wireless situated service, it was deemed necessary to create a situation within which there was a true risk involved in using that technology. To create this risk, a wireless Internet service '*Fastnet*' was created by the researcher and made available to members of the public in two real-world cafes. People who encountered the *Fastnet* service were free to use the service, but in order to do so would need to supply it with some personal information - specifically their mobile phone number. To measure the impact of linkage-based evidence upon trust-investment behaviour, a physical-linkage cue was embedded within the service connection interface (a webpage) that related to the physical surroundings of the user to a variety of degrees. To achieve this, a typical password based network connection protocol was modified to include physically-linked evidence in the form of a photograph.

As a between-groups experiment, some people who encountered the Fastnet webpage would be presented with a photograph that was taken near to where they were sitting (a '*locative*' image). The remaining participants would be presented with a photograph that, while representing a *location*, that location could be interpreted in two different ways: *Could be here, but could equally be anywhere* (an '*a-locative*' image) or *somewhere, but definitely not here* (an '*anti-locative*' image).

The results of the experiment were found to support only the secondary hypothesis of the design. Those participants who were exposed to an *anti-locative* cue were found to be less likely to trust the service than those who were exposed to an *a-locative* image. In more simple terms, when presented with an interface that contained material that was clearly *not* representative of where they physically were, users would trust that interface less than if it contained material that, while location-based, was more generic. Though the results of the experiment were unusual, the experiment was able to show that even simple attempts at providing linkage were able to positively modify user behaviour in response to what was actually a potentially malicious wireless service. However, as a consequence of needing to conduct the experiment without the experimenter himself being physically present during the data collection process, this methodological practice was found to have both positive and negative aspects. People who attempted to use the *Fastnet* service were not aware that they were engaging in an experiment. This was considered as being crucial to creating a sense that there were risks involved in engaging with *Fastnet* and that any degree of risk that users considered as being present would not contaminated through the known presence of a researcher. However, maintaining the deception placed severe constraints on the amount of data that could be captured from the experiment. Without the experimenter present, the methodology rendered access to participant reasoning as to the choices they made during their encounter with *Fastnet* unavailable.

8.3 Contributions of the thesis

8.3.1 Immediate contributions

The first two contributions of the thesis relate to two aspects of human-computer trust in pervasive computing. The first aspect was the development of a better understanding of human-computer trust as it relates to nascent pervasive technologies. The second aspect was the development of evidential cues of service trustworthiness that are relevant to the characteristics of pervasive technologies. In addition, the development and implementation of an *unattended* field study (reported in chapter seven) forms a useful third contribution, containing as it does useful experiential information to practitioners working in similar fields. A discussion of this third contribution can be found later in this chapter (*conducting field research into trust: ethical and practical considerations for researchers*).

Regarding the first two contributions, at the time the research was conducted (2006-2009) interactive technology users, though found to be experienced in a range of risk-relevant activities using their own equipment at home, were also found as having substantially less experience conducting similar experiences in public spaces. this was particularly the case when using non-personal equipment. However, regarding the technologies that they currently owned and used, users considered their privacy and security to be very important; they engaged in a number of practices to support the protection of this information, and demonstrated a good working knowledge of a range of known computer-mediated threats. The finding of chapter seven however, that a third of people who encountered a public-access wireless networking service for the first time would trust that service with personal information indicates that providers of such services should consider protecting their users. In a public wireless Internet service use scenario, users were found to be generally willing to accept a degree of vulnerability that their personal information could be compromised, but only when they considered their immediate need to use such a technology would warrant that risk. However, in order to engage with public wireless technology services for riskier endeavours such as secure email account access and financial transactions, users tended to require substantial reassurance that the services they encountered were both trustworthy, genuine and secure. However, a gulf was also found to exist between what users perceive the security value of a service to be and what that security value actually was.

Practical methods by which pervasive situated service designers can present their products as being both trustworthy and *actually* secure formed the second main contribution of the thesis. Through a process of development and evaluation, the thesis was able to describe a means by which the trustworthiness of benevolent urban pervasive services could be conveyed to users through providing user-verifiable links between a wireless service and the physical environment within which it was encountered. This technique was termed 'linkage'. Importantly, the use of this information also allowed for the discrimination of trustworthy services from similar (and possibly malevolent) services, thus providing users with confidence that their trust, if invested, would likely be well-placed. Across a number of experimental studies, various techniques for providing physical- digital-world linkage were posited and evaluated. Physical-linkage cues were found to increase user trust in an associated service enough for participants to accept the risks of accessing their personal email account. By including virtual-linkage based evidence also, user confidence was found to be increased further still,

to a level where they would consider trusting a service enough to perform even riskier (e.g. personal finance based) activities.

8.3.2 Lasting contributions

That technologies develop and change over time is a continuing challenge to researchers keen to make lasting contributions in an ever changing world. With respect to pervasive computing, the pace of development since this research was originally conceived has increased substantially.

Gaining a secure connection to the Internet from mobile devices, particularly mobile phones is now commonplace. The services offered by cellular network providers are now able to offer an end-to-end secure Internet connection experience that can at least match, if not exceed the experience previously only offered using more dedicated protocols such as WiFi. Concurrently, persistently Internet-connected 'smart' phones and tablet devices have penetrated the public consciousness to the degree where to *not* have such technology within immediate reach can almost be considered unusual.

Bearing these developments in mind, it is perhaps easy to dismiss some of the contributions of the research reported here as being out of date. However, it is the author's belief that the production of this work is actually rather timely. In 2010, the world described in the introduction to this thesis is now generally accepted as an everyday reality. Thus, while it is true that the specific type of pervasive service examined here (public access Internet gateways) may conceivably become less rather than more prevalent in public life, other services that involve ad-hoc connections between what is becoming a highly varied ecology of mobile devices will surely continue to rise. What will certainly not change is the fact that personal / private information will always be of value to opportunists. While the types of service that a pervasive computing world can offer will change, many of the threats and risks attached to their use will not.

8.4 Limitations of the thesis: Advice and recommendations for researchers

8.4.1 Developing the theoretical foundation of the thesis

Viewed in hindsight, it is worth noting that the concept of *linkage*, upon which a substantial amount of the empirical work reported here were developed around, was itself developed organically over the course of the research project. Given the limited timeframe within which to produce a coherent thesis, allowing this process to occur concurrently with the development of an ongoing research programme is risky. With respect to the research reported here, several decision points were encountered along the way as to which path the research should take. A particular case in point was the decision to explore the role of people present

in the immediate physical environment (referred to as 'social linkage'). Upon completion of the lab study reported in chapter five, within which two forms of linkage were identified and evaluated, a decision needed to be made as to how to take the linkage concept forward. Put simply, the choice was to either:

1. Explore in more depth the role of the two forms of linkage as identified in the studies already conducted (i.e. *physical* and *virtual* linkage).
2. Explore the potential for other aspects of the environment to useful in how people considered linkage.

The decision to explore what later became *social linkage* was based upon Kindberg et al's (2002) [72] observation that the physical world can be considered as being made up of three general categories of entity: *people*, *places* and *things*. While the investigation carried out in chapter five considered *places* and *things*, it said little, if anything, about *other people* present. Given a short amount of time with which to conduct the research, the author felt it more appropriate to move towards a perhaps more complete model of linkage by examining the role of people as a highly salient artefact within public space. A consequence of this choice was that the depth of analysis dedicated to physical / virtual linkage would likely be compromised.

With the benefit of hindsight, it is reasonable to suggest that the research would have benefitted from more time spent planning the entire empirical research component of the thesis prior to completing any individual experiment. In some instances, particularly with the *social linkage* study, the consequences of a sudden change of direction might easily have been avoided. However, the nature of conducting individual research within a larger team (as was the case here) can often place additional strain on the time available to conduct pre-planning. Researchers working in a similar environment should be mindful of this, particularly with respect to less experienced researchers who are concurrently learning their craft.

8.4.2 Conducting field research into trust

Ethical implications of empirical trust research: As has been mentioned frequently throughout this thesis, measuring the emergence of trust related attitudes and behaviours is difficult, requiring an uncommon degree of subtlety and care. Designing experiments around trust will almost certainly involve some degree of deception on the part of the researcher as to the true nature of the research he/she is conducting. This issue is made particularly salient with respect to investigations of *trust investment* behaviour, which must involve a plausible (if in reality, bogus) degree of risk before trust behaviours can be said to emerge. However, even with experiments designed to explore a participant's *intention-to-trust*, the author found that, by allowing the participant to elicit discussions about trust as a natural response to a potentially risky situation, a deeper understanding of a participants reasoning about their trust / distrust of a particular entity or situation could be obtained.

Whenever a researcher employs deception within the design of their empirical studies, there are immediate ethical implications which cannot be ignored. All of the experiments reported in this thesis utilised deception to some degree, hence each design was subject to a number of revisions before they were introduced to 'live' participants. Throughout the design process, each experimental design had to be vetted and approved by the University of Bath department of Psychology ethics board. This process is time consuming and often frustrating. However, it is worth reiterating to other researchers who come to read this work that the role of a University Ethics board protects the researcher as well as the participant; in the event that a complaint is raised, the University is in a position to support the researcher, having previously approved the research.

Ecological validity: A recurrent concern of academic researchers is the degree to which results derived from laboratory study can be considered ecologically valid. With research conducted into psychological concepts, especially trust, the choice between using a laboratory-based or field-based experimental design is particularly salient and thus worthy of further discussion.

Within the research reported in this thesis, the author felt strongly that, in order to gain the most useful insight into the mechanisms involved in human-computer trust behaviour, particularly with regard to how risks are identified and evaluated, he would benefit greatly by conducting his experiments within public spaces. Conducting research in public space is, however, a difficult process for a number of reasons. It requires that the researcher both identify, evaluate and potentially control a myriad of factors that are generally not present in the more sterile laboratory environment. It also demands a great deal of time and tenacity; the researcher must often obtain official permission to work within a public space. As an example of this, the *WiFi phishing* study reported in chapter seven required many months of preparatory negotiation work with the University and the owners of various publicly-accessible spaces before it could be made 'active' in the field.

All of the experiments reported in this thesis were conducted in public spaces. Two involved the use of real-world cafes, and one involved the use of a University Library / social space. In each case, the author needed to identify the key people who had control over those spaces, introduce them to the nature of the research and convince them to give their consent for the space to be used. It was fortuitous that a project partner involved in the larger *Cityware* project had a pre-existing relationship with the Bristol-based cafe used as one of the venues in chapter seven. The venue had supported several previous research projects in the past, and the management of the cafe were both aware of the academic process and sympathetic to the goals of the research. However, the management of the venue were still (at least initially) very uneasy about hosting an experiment that was designed to intentionally deceive and potentially inconvenience their customers. That a well-developed relationship between the cafe and the University existed allowed the researcher time to address their concerns; in retrospect, it would be likely that the use of such a venue would not have been granted had there been no prior relationship from which to draw.

In the case of the 'WiFi phishing' experiment reported in chapter seven, to allow the experiment to proceed, the author worked closely with the management to address their concerns directly. Drafts of the experimental design were submitted for their approval regularly and their concerns were included during the ethical review process conducted by the University

of Bath. Consequently, several aspects of the original design were changed to minimise the potential ethical impact of the experiment upon members of the public. The most direct example of this is in the mechanism by which risk was created within the experiment. The original design aimed to utilise the credit card numbers of participants who would not then be aware that they were participating in a research experiment. Though substantial efforts were made to ensure that these numbers would never be put at risk of theft (either through their storage or transmission), it was felt that the potential for participant distress was too high. The eventual compromise, as reported in chapter seven, was to use the participant's mobile phone number. However, even at this lesser level of potential risk, significant efforts were made to protect this data - well above and beyond what might be considered appropriate in an otherwise analogous laboratory based study.

Researchers who intend to conduct studies like the ones described in this thesis should be mindful of the substantially increased timescales involved in conducting experiments in public space. This needs to be considered carefully before abandoning the laboratory. Further, researchers should seek wherever possible to identify and develop working relationships with the key people in control of potential study spaces, preferably well in advance of the development of a final experimental design. Key stakeholders should be identified early and managed directly. Finally, the researcher should take care to ensure that the true nature of the research be presented in as transparent a way as possible throughout the development process.

Measuring trust investment behaviour while also capturing participant reasoning: Within the thesis, a number of research methodologies were employed to investigate the role of human-computer initial-situational trust in pervasive situated services. This mixed-methods research plan offered a number of beneficial aspects, but also a number of negative aspects. As was found over the course of the research programme reported in this thesis, *intention to trust* and the actual *investiture of trust* are indeed very different things. Developing an empirical design that could effectively measure both simultaneously proved to be an extremely difficult task. As such, the experiments that were conducted in this thesis were only able to measure one or the other. In the experiment that measured trust-investment behaviour (chapter seven: *WiFi phishing*) for example, it was strongly felt that the presence of a researcher during the experiment would fatally compromise the ability for that experiment to capture true trust investment behaviour. Whether this was indeed the case in actuality cannot be known, and this issue was a serious weakness of the protocol used in the research reported in that chapter. The unattended nature of the experiment, while offering strength insofar as its ability to provide a means by which true trust investment could be observed and recorded, was weakened by its subsequent inability to capture participant reasoning in-situ. The development of an experimental device that can facilitate both forms of data collection (thus providing for the capture of both trust-investment behaviour and the reasoning behind it), remains the subject of ongoing research by the author.

8.5 Future work

8.5.1 Future development of the linkage concept

A significant issue of the research presented in this thesis was related to the way in which the concept of *linkage*, and its three (thus far) discovered sub-components came to be 'discovered' and examined. While the concept that linkage could be achieved using physical and virtual means had been discussed elsewhere in the literature, the identification of the *social* form of linkage emerged naturally, and somewhat late in the research process. As previously mentioned (see section *developing the theoretical foundation of the thesis*, this chapter), one effect of this late development was that the other two forms of linkage could not be examined in more depth given the time available. Further work would seek to expand upon the investigations of all three forms of linkage discovered thus far, and would also seek to compare their individual / combined effects upon initial-situational trust in technologies developed for public consumption in the time since this research was completed.

8.5.2 Refining the *Interlock* protocol user experience

As a final contribution, one of the two virtually-linked protocols that was developed during the thesis (*Interlock*) was able to offer *actual* security value as a means of facilitating secure ad-hoc device association in a pervasive computing environment. However, its novelty compared to more established user-system authentication practices such as the entry of a password appeared to also have a negative impact on its *perceived* security value. Refining the interaction, such that this gulf might be reduced, is the subject of ongoing research.

The complexity of the *Interlock* protocol user-authentication process was suspected as being a major reason why some participants failed to fully accept it. In further refinements of *Interlock*, reducing the number of steps involved in completing the interaction (without changing the underlying protocol) would be explored. Further, removing the requirement for a physically-fixed public display would be an ideal advance. By using a mobile phone instead of a public display, two issues that were found as being important to users of the *Interlock* protocol - a lack of privacy and the difficulty in making visual comparisons from a distance - could be immediately addressed. To replace the physical-linkage aspect that would be lost with the removal of the screen, the user could pick up a token from a point that is strongly physically linked to the cafe (in all four senses identified as being important to the participants in that study: attachment, legitimacy, visibility and accessibility). A possible example of this might be a card handed over by a member of staff. Upon the card might be a 2D barcode that is readable with the user's mobile phone, thus linking it securely to the venue and obtaining a means of displaying the content as previously made available via the public screen.

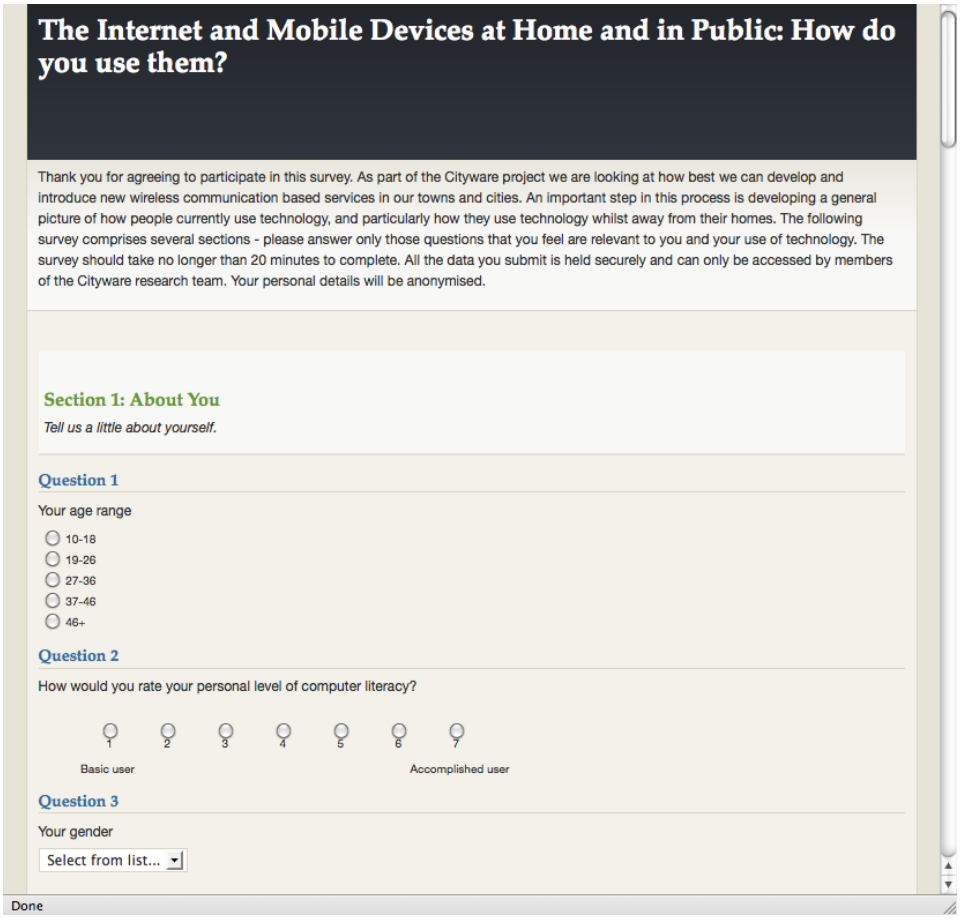
8.6 Conclusion

This thesis sought to uncover the processes and mechanisms by which people assess the trustworthiness of a pervasive situated service upon their first encounter. In this thesis, two types of situation that involved a requirement for human-computer trust and relevancy to pervasive computing service usage were examined across a number of experimental studies. These two situations were 1) service discovery and 2) ad-hoc device association. Results from all of the experiments indicated that, in the absence of prior knowledge, people would seek to identify and use contextual evidence to support their assessment as to the potential trustworthiness of a new situated service. To support this finding, a number of methods of providing salient and credible evidence as to the genuineness of a situated service were explored. The provision of user-verifiable 'links' between a wireless digital service and the physical world in which it is encountered were found to be an effective evidential cue to service trustworthiness. Channels of linkage evidence were found as being able to take a physical, virtual or social form.

Appendix A

Companion to chapter 3

A.1 Survey



The Internet and Mobile Devices at Home and in Public: How do you use them?

Thank you for agreeing to participate in this survey. As part of the Cityware project we are looking at how best we can develop and introduce new wireless communication based services in our towns and cities. An important step in this process is developing a general picture of how people currently use technology, and particularly how they use technology whilst away from their homes. The following survey comprises several sections - please answer only those questions that you feel are relevant to you and your use of technology. The survey should take no longer than 20 minutes to complete. All the data you submit is held securely and can only be accessed by members of the Cityware research team. Your personal details will be anonymised.

Section 1: About You

Tell us a little about yourself.

Question 1

Your age range

- 10-18
- 19-26
- 27-36
- 37-46
- 46+

Question 2

How would you rate your personal level of computer literacy?

1 2 3 4 5 6 7

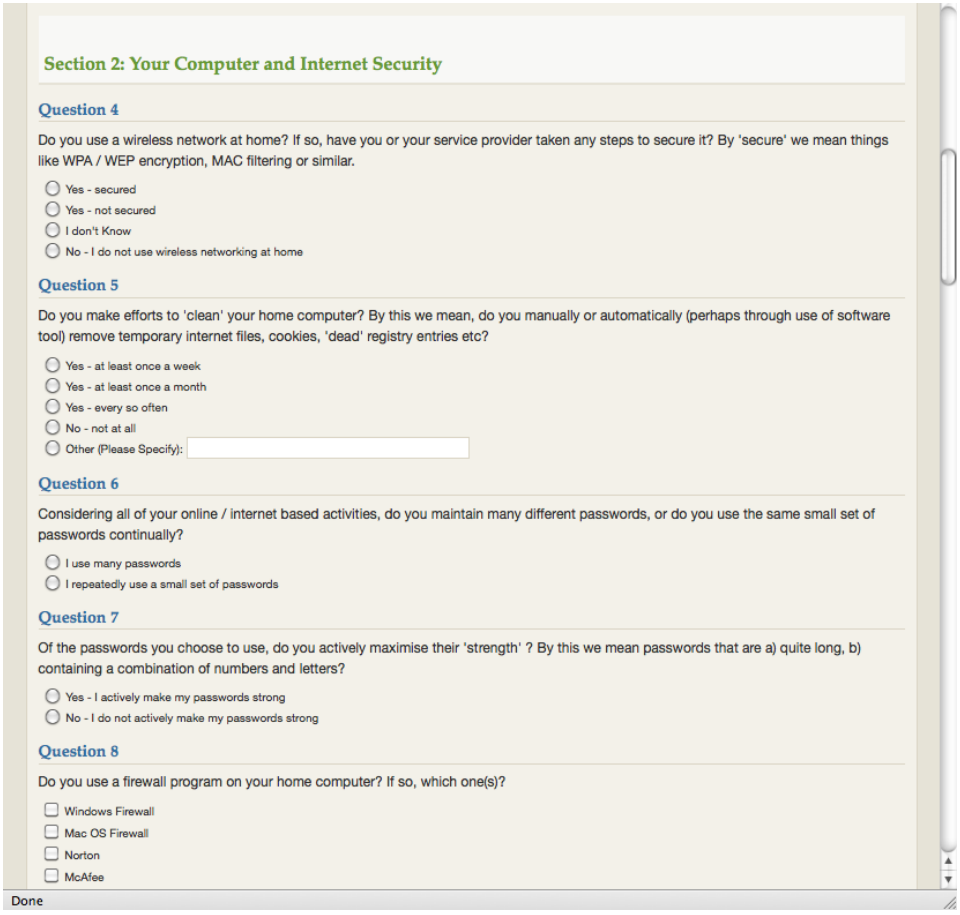
Basic user — Accomplished user

Question 3

Your gender

Select from list...

Done



Section 2: Your Computer and Internet Security

Question 4

Do you use a wireless network at home? If so, have you or your service provider taken any steps to secure it? By 'secure' we mean things like WPA / WEP encryption, MAC filtering or similar.

- ◯ Yes - secured
- ◯ Yes - not secured
- ◯ I don't Know
- ◯ No - I do not use wireless networking at home

Question 5

Do you make efforts to 'clean' your home computer? By this we mean, do you manually or automatically (perhaps through use of software tool) remove temporary internet files, cookies, 'dead' registry entries etc?

- ◯ Yes - at least once a week
- ◯ Yes - at least once a month
- ◯ Yes - every so often
- ◯ No - not at all
- ◯ Other (Please Specify):

Question 6

Considering all of your online / internet based activities, do you maintain many different passwords, or do you use the same small set of passwords continually?

- ◯ I use many passwords
- ◯ I repeatedly use a small set of passwords

Question 7

Of the passwords you choose to use, do you actively maximise their 'strength' ? By this we mean passwords that are a) quite long, b) containing a combination of numbers and letters?

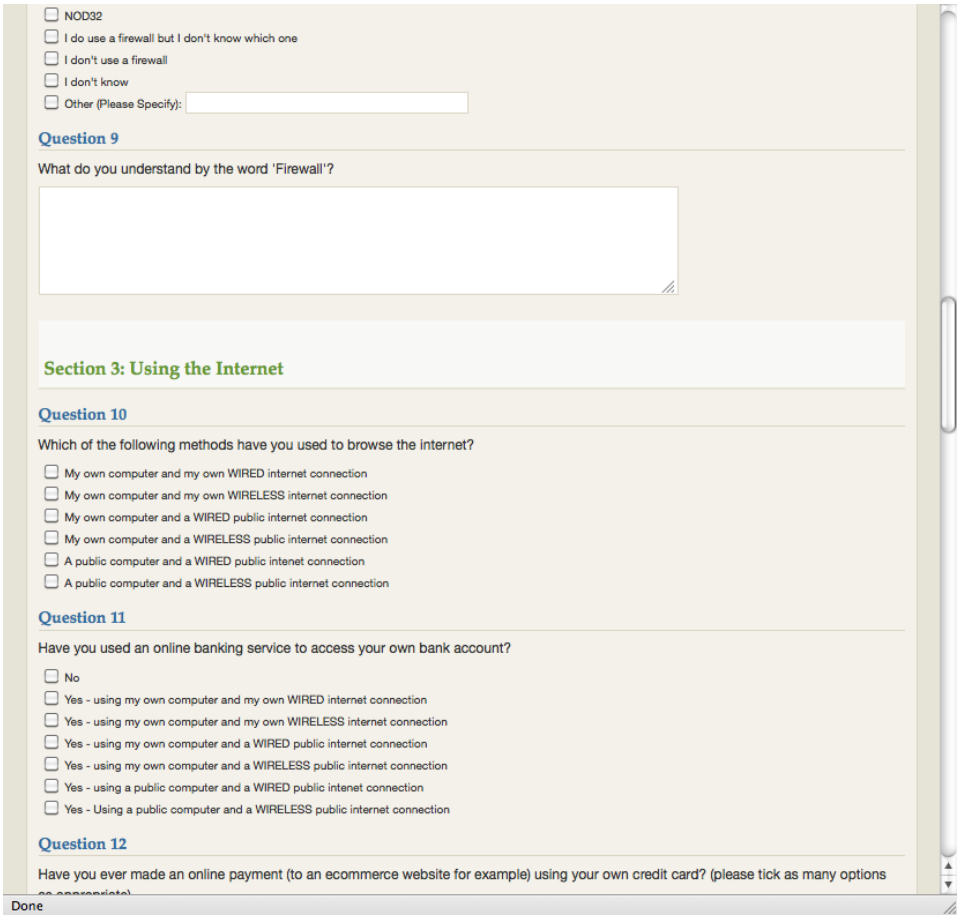
- ◯ Yes - I actively make my passwords strong
- ◯ No - I do not actively make my passwords strong

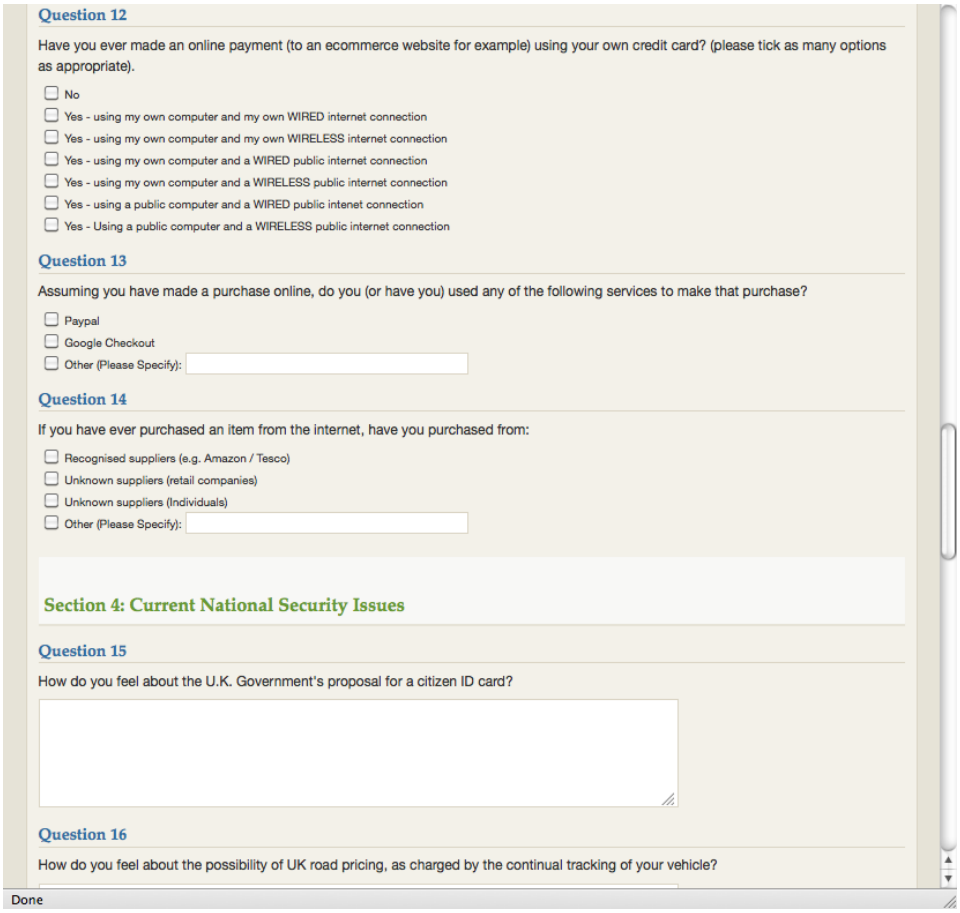
Question 8

Do you use a firewall program on your home computer? If so, which one(s)?

- ☐ Windows Firewall
- ☐ Mac OS Firewall
- ☐ Norton
- ☐ McAfee

Done


NOD32
I do use a firewall but I don't know which one
I don't use a firewall
I don't know
Other (Please Specify):
Question 9
What do you understand by the word 'Firewall'?
Section 3: Using the Internet
Question 10
Which of the following methods have you used to browse the internet?
My own computer and my own WIRED internet connection
My own computer and my own WIRELESS internet connection
My own computer and a WIRED public internet connection
My own computer and a WIRELESS public internet connection
A public computer and a WIRED public intenet connection
A public computer and a WIRELESS public internet connection
Question 11
Have you used an online banking service to access your own bank account?
No
Yes - using my own computer and my own WIRED internet connection
Yes - using my own computer and my own WIRELESS internet connection
Yes - using my own computer and a WIRED public internet connection
Yes - using my own computer and a WIRELESS public internet connection
Yes - using a public computer and a WIRED public intenet connection
Yes - Using a public computer and a WIRELESS public internet connection
Question 12
Have you ever made an online payment (to an ecommerce website for example) using your own credit card? (please tick as many options
Done



Question 12

Have you ever made an online payment (to an ecommerce website for example) using your own credit card? (please tick as many options as appropriate).

- [ ] No
- [ ] Yes - using my own computer and my own WIRED internet connection
- [ ] Yes - using my own computer and my own WIRELESS internet connection
- [ ] Yes - using my own computer and a WIRED public internet connection
- [ ] Yes - using my own computer and a WIRELESS public internet connection
- [ ] Yes - using a public computer and a WIRED public intenet connection
- [ ] Yes - Using a public computer and a WIRELESS public internet connection

Question 13

Assuming you have made a purchase online, do you (or have you) used any of the following services to make that purchase?

- [ ] Paypal
- [ ] Google Checkout
- [ ] Other (Please Specify):

Question 14

If you have ever purchased an item from the internet, have you purchased from:

- [ ] Recognised suppliers (e.g. Amazon / Tesco)
- [ ] Unknown suppliers (retail companies)
- [ ] Unknown suppliers (Individuals)
- [ ] Other (Please Specify):

Section 4: Current National Security Issues

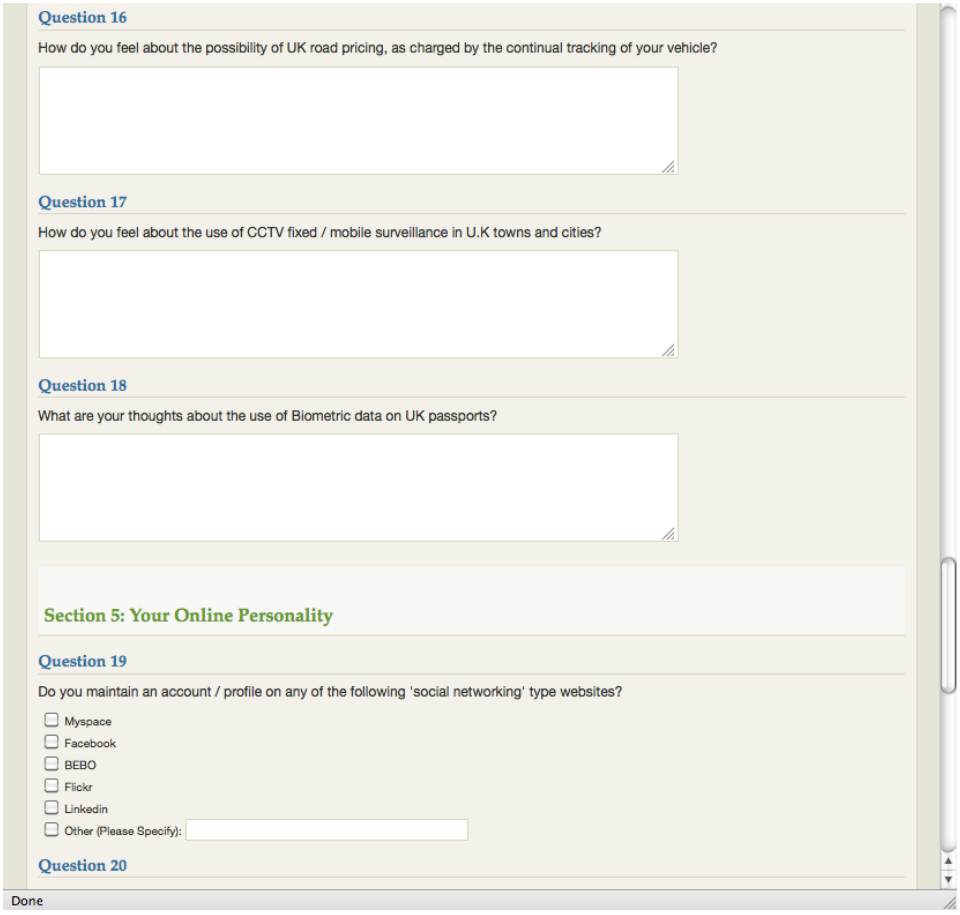
Question 15

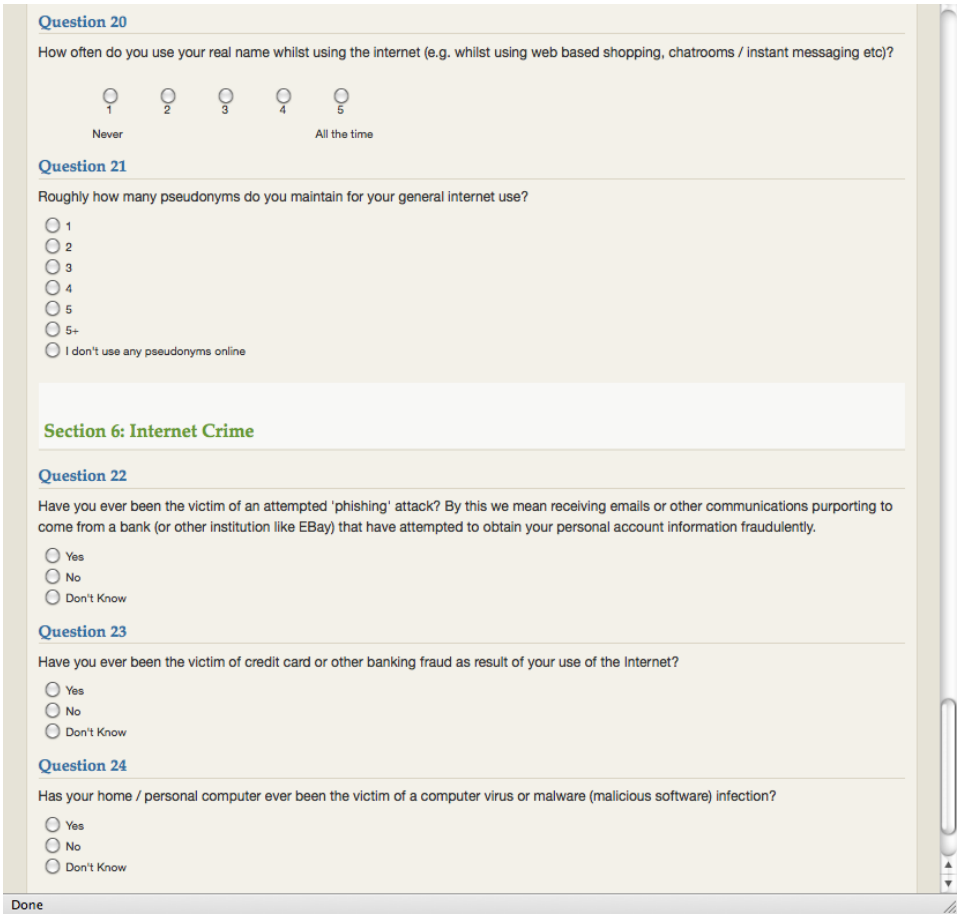
How do you feel about the U.K. Government's proposal for a citizen ID card?

Question 16

How do you feel about the possibility of UK road pricing, as charged by the continual tracking of your vehicle?

Done


Question 16
How do you feel about the possibility of UK road pricing, as charged by the continual tracking of your vehicle?
Question 17
How do you feel about the use of CCTV fixed / mobile surveillance in U.K towns and cities?
Question 18
What are your thoughts about the use of Biometric data on UK passports?
Section 5: Your Online Personality
Question 19
Do you maintain an account / profile on any of the following 'social networking' type websites?
Myspace
Facebook
BEBO
Flickr
Linkedin
Other (Please Specify):
Question 20
Done



Question 20

How often do you use your real name whilst using the internet (e.g. whilst using web based shopping, chatrooms / instant messaging etc)?

1 2 3 4 5

Never All the time

Question 21

Roughly how many pseudonyms do you maintain for your general internet use?

- 1
- 2
- 3
- 4
- 5
- 5+
- I don't use any pseudonyms online

Section 6: Internet Crime

Question 22

Have you ever been the victim of an attempted 'phishing' attack? By this we mean receiving emails or other communications purporting to come from a bank (or other institution like EBay) that have attempted to obtain your personal account information fraudulently.

- Yes
- No
- Don't Know

Question 23

Have you ever been the victim of credit card or other banking fraud as result of your use of the Internet?

- Yes
- No
- Don't Know

Question 24

Has your home / personal computer ever been the victim of a computer virus or malware (malicious software) infection?

- Yes
- No
- Don't Know

Done

Appendix B

Companion to chapter 5

B.1 *Bertorelli's* experiment: Consent form

B.2 *Bertorelli's* experiment: Instructions for participants

cityware
urban design & pervasive systems

UNIVERSITY OF BATH

Participant Consent

Please read the following information carefully. The information in this consent form is provided so that you can decide whether you wish to participate in our study. It is important that you understand that your participation is considered voluntary. This means that even if you agree to participate you are free to withdraw from the study at any time should you so wish.

During the study you will be videotaped so that we can observe you as you complete the task. These videotapes will be stored securely and viewed only by project members, after which they will be destroyed. All information, visual, auditory or otherwise supplied by you will be entirely confidential, and all data will be made anonymous before any analysis of the data takes place.

Some Information about You

Name: _____

Gender: _____

Age Range (please circle): 18-25 26-30 31-45 45+

Occupation: _____

By providing my signature, I agree to participate in this study

Signature: _____

Print name: _____

cityware
urban design & pervasive systems

UNIVERSITY OF BATH

Thank you for agreeing to take part in this study. Your help is invaluable to us.

Imagine you are visiting a café to try to get an Internet connection. You have your laptop computer with you, and you know that Bertorelli's cafe happens to have a wireless Internet access point that you can use during your visit.

When you sit down at a table and try to connect to the wireless network you see that there are six different wireless networks that *appear to be* provided by Bertorelli's café. You know that at least one of these is genuinely provided by Bertorelli's cafe, but one or more may be from another source pretending to be Bertorelli's cafe. There are two ways in which wireless connections are susceptible to attack:

- Somebody might 'listen-in' to the wireless communications made between your laptop and a genuine access point to the Internet such as the one provided by Bertorelli's café.
- Somebody might have created an entirely fake network to which you can still connect, potentially giving away passwords or other information and inadvertently giving the faker access to your computer.

What we would like you to do…

1. **Please attempt to connect to the Internet using each of the six wireless networks, in the order you will be given**. Each wireless network has a different means of establishing a connection to the Internet. You will be instructed when each connection attempt has been completed.
2. We are very interested to see how people make decisions about the different access methods and how they decide which are more likely to be genuine than others. **While you are working on the task, we would like you to talk through what you are thinking**. An experimenter will sit with you as you complete the task.
3. Once you have completed the task, we will ask you some questions about which of the wireless networks you think are most likely to have been genuinely provided by Bertorelli's café.

Appendix C

Companion to chapter 6

C.1 *Social Linkage* experiment: Consent form

C.2 *Social Linkage* experiment: Instructions for participants

cityware
urban design & pervasive systems

UNIVERSITY OF BATH

Participant Consent

Please read the following information carefully. The information in this consent form is provided so that you can decide whether you wish to participate in our study. It is important that you understand that your participation is voluntary. This means that even if you agree to participate you are free to withdraw from the study at any time should you so wish.

During the study your comments will be recorded as digital audio files. These audio files will be stored securely and reviewed only by project members. All information, written, auditory or otherwise supplied by you will be entirely confidential. In the event that data generated by the study is published, all data will be made anonymous. The lead investigator in this study is Chris Bevan, and he can be contacted at crb23@bath.ac.uk. For more information about this research, please visit http://www.cityware.org.uk.

Some Information about You

Name: _____

Gender: _____

Age Range (please circle): 18-25 26-30 31-45 45+

Occupation: _____

By providing my signature, I agree to participate in this study and understand that I may withdraw my consent at any time

Signature: _____

Date: _____

cityware
urban design & pervasive systems

UNIVERSITY OF BATH

Imagine you are visiting a public library to try to get an Internet connection. You have your laptop computer with you, and you know that this library happens to have a wireless Internet access point that you can use during your visit.

When you sit down at a table and try to connect to the wireless network you see that there are in fact two different wireless networks that both *appear to be* provided by the library. **You know that the library genuinely provides one of these networks. The other network may be from another source pretending to be the library and should be considered as potentially fake.** You are aware that wireless networks are vulnerable to attack; somebody might have created an entirely fake network to which you can still connect, potentially giving away passwords or other information and inadvertently giving the faker access to your computer.

The following point is very important!!

Please consider all information you encounter through the Wireless Network Connection interface used during the study as being *live* and *current.*

What we would like you to do…

1. **Please attempt to connect to the Internet using our system (you will be instructed how to do this).** You will perform this task several times and you will be instructed when each connection attempt has been completed.
2. You can find the available networks by clicking the button labelled '*scan for wireless networks*'. **Please choose the network you consider to be most likely to be authentic** and click the button labelled "*connect".*
3. When you click "connect", the system will ask you to **rate your choice based on how confident you are about the authenticity of that particular network.**
4. We are interested to see how people make decisions about the different access methods and how they decide which are more likely to be genuine than others. **While you are working on the task, we would like you to talk through what you are thinking.** An experimenter will sit with you as you complete the task.
5. Once you have completed the tasks, we will ask you some questions about how you came to decide which of the networks was genuine and which of the networks was fake.

Appendix D

Companion to chapter 7

D.1 Experimental materials: Photos used in a ranking exercise to generate the a-locative and locative image conditions

(a) Local (Bristol) Image 1

(b) Local (Bristol) Image 2

(c) Local (Bristol) Image 3

Figure D.1: Image ranking exercise, Local image conditions. All photos were taken in the immediate area outside the cafe used in Bristol, U.K.

(a) Generic Image 1

(b) Generic Image 2

(c) Generic Image 3

(d) *Wildcard* Image

Figure D.2: Image ranking exercise, generic image conditions and wildcard. All photos sourced from stock photography.

Bibliography

[1] Abdul-Rahman, A., and Hailes, S. Supporting trust in virtual communities. In *HICSS '00: Proceedings of the 33rd Hawaii International Conference on System Sciences-Volume 6* (Washington, DC, USA, 2000), IEEE Computer Society, p. 6007.

[2] Adams, A., and Sasse, M. A. Users are not the enemy. *Commun. ACM 42*, 12 (1999), 40–46.

[3] Adams, A., and Sasse, M. A. Privacy in multimedia communications: Protecting users, not just data. In *In* (2001), Springer, pp. 49–64.

[4] Adams, J. *Risk*. University College London Press, London, 1995.

[5] Adolphs, R. Recognizing Emotion from Facial Expressions: Psychological and Neurological Mechanisms. *Behavioral and Cognitive Neuroscience Reviews 1*, 1 (2002), 21–62.

[6] Akerlof, G. A. Loyalty filters. *American Economic Review 73* (March 1983), 54–63.

[7] Allport, G. *Pattern and Growth in Personality*. New York: Holt, Rinehart and Winston, 1961.

[8] Axelrod, R. *The Evolution of Cooperation*. Basic Books, New York, 1984.

[9] Balfanz, D., Durfee, G., Grinter, R. E., Smetters, D. K., and Stewart, P. Network-in-a-box: how to set up a secure wireless network in under a minute. In *SSYM'04: Proceedings of the 13th conference on USENIX Security Symposium* (Berkeley, CA, USA, 2004), USENIX Association, pp. 15–15.

[10] Barkow, J. H., Cosmides, L., and Tooby, J. *The Adapted Mind: Evolutionary Psychology and the Generation of Culture.* New York/Oxford: Oxford University Press, 1992.

[11] Barton, J., and Kindberg, T. The challenges and opportunities of integrating the physical world and networked systems, 2001.

[12] Beckwith, R. Designing for ubiquity: The perception of privacy. *IEEE Pervasive Computing 2* (2003), 40–46.

[13] Benford, S., Flintham, M., Drozd, A., Anastasi, R., Rowland, D., Tandavanitj, N., Adams, M., Row-Farr, J., Oldroyd, A., and Sutton, J. Uncle roy all around you: Implicating the city in a location-based performance. In *Advances in Computer Entertainment (ACE 2004)* (2004), ACM Press.

[14] Berg, J., Dickhaut, J., and McCabe, K. Trust, reciprocity, and social history. *Games and Economic Behavior 10*, 1 (1995), 122 – 142.

[15] Boone, R. T., and Buck, R. Emotional expressivity and trustworthiness: The role of nonverbal behavior in the evolution of cooperation. *Journal of Nonverbal Behavior 27* (2003), 163–182. 10.1023/A:1025341931128.

[16] Boyd, J. In community we trust: Online security communication at ebay. *Journal of Computer-Mediated Communication 7*, 3 (2002).

[17] Brenkert, G. G. Trust, morality and international business. *Business Ethics Quarterly 8*, 2 (1998), 293–317.

[18] Brewer, M., and Silver, M. Ingroup bias as a function of task characteristics. *European Journal of Social Psychology 8*, 3 (1978), 393–400.

[19] Butler, M., Leuschel, M., Presti, S. L., and Allsopp, D. Towards a trust analysis framework for pervasive computing scenarios, 2003.

[20] Cassell, J., and Bickmore, T. External manifestations of trustworthiness in the interface. *Commun. ACM 43*, 12 (2000), 50–56.

[21] Castelli, G., Rosi, A., Mamei, M., Zambonelli, F., Università, D., Reggio, M., Via, E., Reggio, A., and Italy, E. A simple model and infrastructure for context-aware browsing of the world. In *In Proceeding of PERCOM '07, IEEE Computer Society* (2007), pp. 229–238.

[22] Catell, R. B. *The Scientific Analysis of Personality*. Baltimore: Penguin Books, 1965.

[23] Chakraborty, D., Joshi, A., Yesha, Y., and Finin, T. Toward distributed service discovery in pervasive computing environments. *IEEE Transactions on Mobile Computing 5* (2006), 97–112.

[24] Chen, Y.-H., and Barnes, S. Initial trust and online buyer behaviour. *Industrial Management Data Systems 107*, 1 (2007), 21–36.

[25] Cheverst, K., Clarke, K., Dewsbury, G., Hemmings, T., Kember, S., Rodden, T., and Rouncefield, M. Designing assistive technologies for medication regimes in care settings. *Universal Access in the Information Society 2* (2003), 235–242. 10.1007/s10209-003-0055-9.

[26] Cialdini, R. B. *Influence: Science and Practice (4th Edition)*. Allyn & Bacon, June 2000.

[27] Corritore, C. L., Kracher, B., and Wiedenbeck, S. On-line trust: concepts, evolving themes, a model. *International Journal of Human-Computer Studies 58*, 6 (2003), 737 – 758. Trust and Technology.

[28] Crabtree, A., Benford, S., Capra, M., Flintham, M., Drozd, A., Tandavanitj, N., Adams, M., and Row Farr, J. The cooperative work of gaming: Orchestrating a mobile sms game. *Computer Supported Cooperative Work (CSCW) 16* (2007), 167–198. 10.1007/s10606-007-9048-1.

[29] Crabtree, A., Benford, S., Rodden, T., Greenhalgh, C., Flintham, M., Anastasi, R., Drozd, A., Adams, M., Row-Farr, J., Tandavanitj, N., and Steed, A. Orchestrating a mixed reality game 'on the ground'. In *CHI '04: Proceedings of the SIGCHI conference on Human factors in computing systems* (New York, NY, USA, 2004), ACM, pp. 391–398.

[30] Dasgupta, P. Trust as a commodity, 2000.

[31] DeBruine, L. M. Facial resemblance enhances trust. *Proceedings of the Royal Society of London. Series B: Biological Sciences 269*, 1498 (2002), 1307–1312.

[32] Delhey, J., and Newton, K. Who trusts? the origins of social trust in seven societies.

[33] Deriaz, M. What is trust? my own point of view.

[34] Deutsch, M. Trust and suspicion. *The Journal of Conflict Resolution 2*, 4 (1958), 265–279.

[35] Deutsch, M. *The Resolution of Conflict: Constructive and Destructive Processes.* Yale University Press, New Haven, 1974.

[36] Dhamija, R. Hash visualization in user authentication. In *CHI '00 extended abstracts on Human factors in computing systems* (New York, NY, USA, 2000), ACM, pp. 279–280.

[37] Dhamija, R., and Tygar, J. D. The battle against phishing: Dynamic security skins. In *SOUPS '05: Proceedings of the 2005 symposium on Usable privacy and security* (New York, NY, USA, 2005), ACM, pp. 77–88.

[38] Dhamija, R., Tygar, J. D., and Hearst, M. Why phishing works. In *CHI '06: Proceedings of the SIGCHI conference on Human Factors in computing systems* (New York, NY, USA, 2006), ACM, pp. 581–590.

[39] Doney, P., and Cannon, J. An examination of the nature of trust in buyer-seller relationships. *Journal of Marketing 61* (April 1997), 35–51.

[40] Downs, J. S., Holbrook, M. B., and Cranor, L. F. Decision strategies and susceptibility to phishing. In *SOUPS '06: Proceedings of the second symposium on Usable privacy and security* (New York, NY, USA, 2006), ACM, pp. 79–90.

[41] Egger, F. N. Affective design of e-commerce user interfaces: How to maximise perceived trustworthiness. In *Proc. Intl. Conf. Affective Human Factors Design* (2001), pp. 317–324.

[42] Erikson, E. H. *Childhood and Society.* New York: Norton, 1950.

[43] Falcone, R., and Castelfranchi, C. Social trust: a cognitive approach. 55–90.

[44] Fisman, R., and Khanna, T. Is trust a historical residue? information flows and trust levels. *Journal of Economic Behavior Organization 38*, 1 (1999), 79 – 92.

[45] Florencio, D., and Herley, C. A large-scale study of web password habits. In *WWW '07: Proceedings of the 16th international conference on World Wide Web* (New York, NY, USA, 2007), ACM, pp. 657–666.

[46] Fogg, B., Marshall, J., Kameda, T., Solomon, J., Rangnekar, A., Boyd, J., and Brown, B. Web credibility research: a method for online experiments and early study results. In *CHI '01: CHI '01 extended abstracts on Human factors in computing systems* (New York, NY, USA, 2001), ACM, pp. 295–296.

[47] Fogg, B. J., Soohoo, C., Danielson, D. R., Marable, L., Stanford, J., and Tauber, E. R. How do users evaluate the credibility of web sites?: a study with over 2,500 participants. In *DUX '03: Proceedings of the 2003 conference on Designing for user experiences* (New York, NY, USA, 2003), ACM, pp. 1–15.

[48] Frank, R. H. If homo economicus could choose his own utility function, would he want one with a conscience? *The American Economic Review 77*, 4 (September 1987), 593–604.

[49] Fukuyama, F. *Trust: The Social Virtues and The Creation of Prosperity.* London: Penguin, 1995.

[50] Gambetta, D. Can we trust trust? In *Trust: Making and Breaking Cooperative Relations* (1988), Basil Blackwell, pp. 213–237.

[51] Giddens, A. *The consequences of modernity.* Stanford: Stanford University Press, 1990.

[52] Giddens, A. *Modernity and Self-Identity.* Cambridge: Polity Press, 1991.

[53] Glaeser, E. L., Laibson, D. I., Scheinkman, J. A., and Soutter, C. L. Measuring trust. *The Quarterly Journal of Economics 115*, 3 (August 2000), 811–846.

[54] Good, D. Individuals, interpersonal relations, and trust, 2000.

[55] Gouldner, A. W. The norm of reciprocity: A preliminary statement. *American Sociological Review 25*, 2 (1960), 161–178.

[56] Grawemeyer, B., and Johnson, H. How secure is your password? towards modelling human password creation. In *First Trust Economics Workshop, University College London, England 23 June 2009* (2009), Available at: http://www.trust-economics.org/TEWorkshopProceedings.pdf.

[57] Grazioli, S., and Jarvenpaa, S. Perils of internet fraud: an empirical investigation of deception and trust with experienced internet consumers. *IEEE Transactions on Systems, Man, and Cybernetics 30*, 4 (2000), 395 – 410.

[58] Greenfield, A. *Everyware: The Dawning Age of Ubiquitous Computing.* New Riders Publishing, Berkeley, CA, 2006.

[59] Hampton-Sosa, W., and Koufaris, M. The effect of web site perceptions on initial trust in the owner company. *Int. J. Electron. Commerce 10*, 1 (2005), 55–81.

[60] Hardin, R. The street-level epistemology of trust. *Politics and Society 21*, 4 (1993), 505.

[61] Hertzum, M., Andersen, H. H. K., Andersen, V., and Hansen, C. B. Trust in information sources: seeking information from people, documents, and virtual agents. *Interacting with Computers 14*, 5 (2002), 575 – 599.

[62] Hofstede, G. *Culture's Consequences*. Beverley Hills, CA: Sage, 1980.

[63] Holmquist, L., Mattern, F., Schiele, B., Alahuhta, P., Beigl5, M., and Gellersen, H.-W. Smart-its friends: A technique for users to easily establish connections between smart artefacts. *Ubicomp 2001: Ubiquitous Computing* (2001), 116–122.

[64] Hull, R., Neaves, P., and Bedford-roberts, J. Towards situated computing. In *In Proceedings of The First International Symposium on Wearable Computers* (1997), pp. 146–153.

[65] Ives, B., Walsh, K. R., and Schneider, H. The domino effect of password reuse. *Commun. ACM 47*, 4 (2004), 75–78.

[66] James, H. S. The trust paradox: a survey of economic inquiries into the nature of trust and trustworthiness. *Journal of Economic Behavior Organization 47*, 3 (2002), 291 – 307.

[67] Jarvenpaa, S. L., and Tractinsky, M. Consumer trust in an internet store: A cross-cultural validation. *Journal of Computer-Mediated Communication* (2000).

[68] Jay, T., and Fraser, D. S. The role of a cohort in the design and evaluation of pervasive systems. In *DIS '08: Proceedings of the 7th ACM conference on Designing interactive systems* (New York, NY, USA, 2008), ACM, pp. 31–39.

[69] Johnson-George, C., and Swap, W. C. Measurement of specific interpersonal trust: Construction and validation of a scale to assess trust in a specific other. *Journal of Personality and Social Psychology 43*, 6 (1982), 1306 – 1317.

[70] Kagal, L., Finin, T., and Joshi, A. Trust-based security in pervasive computing environments. *Computer 34*, 12 (2001), 154–157.

[71] Kim, K. K., and Prabhakar, B. Initial trust and the adoption of b2c e-commerce: The case of internet banking. *SIGMIS Database 35*, 2 (2004), 50–64.

[72] Kindberg, T., Barton, J., Morgan, J., Becker, G., Caswell, D., Debaty, P., Gopal, G., Frid, M., Krishnan, V., Morris, H., Schettino, J., Serra, B., and Spasojevic, M. People, places, things: web presence for the real world. *Mob. Netw. Appl. 7*, 5 (2002), 365–376.

[73] Kindberg, T., Mitchell, J., Grimmett, J., Bevan, C., and O'Neill, E. Authenticating public wireless networks with physical evidence. In *WIMOB '09: Proceedings of the 2009 IEEE International Conference on Wireless and Mobile Computing, Networking and Communications* (Washington, DC, USA, 2009), IEEE Computer Society, pp. 394–399.

[74] Kindberg, T., O'Neill, E., Bevan, C., Kostakos, V., Stanton Fraser, D., and Jay, T. Measuring trust in wi-fi hotspots. In *CHI '08: Proceeding of the twenty-sixth annual SIGCHI conference on Human factors in computing systems* (New York, NY, USA, 2008), ACM, pp. 173–182.

[75] Kindberg, T., Sellen, A., and Geelhoed, E. Security and trust in mobile interactions: A study of users' perceptions and reasoning. In *Proceedings of the International Conference on Ubiquitous Computing (UbiComp) 2004* (2004), vol. 32, pp. 196–213.

[76] Kindberg, T., and Zhang, K. Secure spontaneous device association. *UbiComp 2003: Ubiquitous Computing* (2003///), 124–131.

[77] Kindberg, T., and Zhang, K. Validating and securing spontaneous associations between wireless devices. *Information Security* (2003///), 44–53.

[78] Kobsa, A., Sonawalla, R., Tsudik, G., Uzun, E., and Wang, Y. Serial hook-ups: a comparative usability study of secure device pairing methods. In *SOUPS '09: Proceedings of the 5th Symposium on Usable Privacy and Security* (New York, NY, USA, 2009), ACM, pp. 1–12.

[79] Koufaris, M., and Hampton-Sosa, W. The development of initial trust in an online company by new customers. *Information Management 41*, 3 (2004), 377 – 397.

[80] Kumar, A., Saxena, N., Tsudik, G., and Uzun, E. A comparative study of secure device pairing methods. *Pervasive and Mobile Computing 5*, 6 (2009), 734 – 749. PerCom 2009.

[81] Lee, H. G. Understanding dynamics between initial trust and usage intentions of mobile banking. *Information Systems Journal 19* (May 2009), 283–311(29).

[82] Lee, M., and Turban, E. A trust model for consumer internet shopping. *Int. J. Electron. Commerce 6*, 1 (2001), 75–91.

[83] Lewicki, R., and Bunker, B. Developing and maintaining trust in work relationships. In *Conflict, Cooperation and Justice*, B. Bunker and J. Rubin, Eds. Jossey-Bass, San Francisco, 1995.

[84] Lewis, J. D., and Weigert, A. Trust as a Social Reality. *Social Forces 63*, 4 (1985), 967–985.

[85] Luhmann, N. *Trust and Power*. Chichester: Wiley, 1979.

[86] Macy, M. W., and Skvoretz, J. The evolution of trust and cooperation between strangers: A computational model. *American Sociological Review 63*, 5 (1998), 638–660.

[87] Mahrer, A. The role of expectancy in delayed reinforcement. *I. Exf. Psychol. 62* (1956), 101–105.

[88] Malhotra, D. Trust and reciprocity decisions: The differing perspectives of trustors and trusted parties. *Organizational Behavior and Human Decision Processes 94*, 2 (2004), 61–73.

[89] marc Seigneur, J., and Jensen, C. D. The role of identity in pervasive computational trust. In *Privacy, Security and Trust within the Context of Pervasive Computing* (2005), Springer US, pp. 65–75.

[90] Marschall, M. J., and Stolle, D. Race and the city: Neighborhood context and the development of generalized trust. *Political Behavior 26*, 2 (2004), 125–153.

[91] Mayer, R. C., Davis, J. H., and Schoorman, F. D. An integrative model of organizational trust. *The Academy of Management Review 20*, 3 (1995), 709–734.

[92] Mayerson, D., Weick, K. E., and Kramer, R. M. Swift trust and temporary groups. In *Trust in organizations: frontiers of theory and research*, R. Kramer and T. Tyler, Eds. Sage Publications, 1996, ch. 9.

[93] Mayrhofer, R., and Gellersen, H. Shake well before use: Authentication based on accelerometer data. In *In Pervasive* (2007), Springer, pp. 144–161.

[94] Mayrhofer, R., Gellersen, H., and Hazas, M. Security by spatial reference: Using relative positioning to authenticate devices for spontaneous interaction. *UbiComp 2007: Ubiquitous Computing* (2007), 199–216.

[95] Mayrhofer, R., and Welch, M. A human-verifiable authentication protocol using visible laser light. *Availability, Reliability and Security, International Conference on 0* (2007), 1143–1148.

[96] McAllister, D. J. Affect- and cognition-based trust as foundations for interpersonal cooperation in organizations. *The Academy of Management Journal 38*, 1 (February 1995), 24–59.

[97] McCune, J. M., Perrig, A., and Reiter, M. K. Seeing-is-believing: Using camera phones for human-verifiable authentication. *Security and Privacy, IEEE Symposium on 0* (2005), 110–124.

[98] McKnight, D., Choudhury, V., and Kacmar, C. Developing and validating trust measures for e-commerce: An integrative typology. *Information Systems Research 13*, 4 (2002), 334 – 359.

[99] McKnight, D., Choudhury, V., and Kacmar, C. The impact of initial consumer trust on intentions to transact with a web site: A trust building model. *Journal of Strategic Information Systems 11* (2002), 297 – 323.

[100] McKnight, D. H., and Chervany, N. L. Trust and distrust definitions: One bite at a time. In *Proceedings of the workshop on Deception, Fraud, and Trust in Agent Societies held during the Autonomous Agents Conference* (London, UK, 2001), Springer-Verlag, pp. 27–54.

[101] McKnight, D. H., Cummings, L. L., and Chervany, N. L. Initial trust formation in new organizational relationships. *The Academy of Management Review 23*, 3 (1998), 473–490.

[102] McKnight, D. H., Kacmar, C. J., and Choudhury, V. Shifting factors and the ineffectiveness of third party assurance seals: A two-stage model of initial trust in a web business. *Electronic Markets 14* (September 2004), 252–266(15).

[103] Mischel, W. Preference for delayed reinforcement and social responsibility. *Journal of abnormal social psychology 62* (1961), 1–7.

[104] Misztal, B. A. *Trust in Modern Societies the Search for the Bases of Social Order.* Cambridge: Polity Press, 1998.

[105] Mohtashemi, M., and Mui, L. Evolution of indirect reciprocity by social information: The role of trust and reputation in evolution of altruism, 2003.

[106] Mollering, G. The nature of trust: From georg simmel to a theory of expectation, interpretation and suspension. *Sociology 35*, 02 (2001), 403–420.

[107] Morgernstern, O., and Von Neumann, J. *Theory of games and economic behavior*. Princeton University Press, 1953.

[108] National Institute of Standards and Technology. *FIPS PUB 112: Standard for Password Usage*. May 1985.

[109] Nilsson, M., Adams, A., and Herd, S. Building security and trust in online banking. In *CHI '05: CHI '05 extended abstracts on Human factors in computing systems* (New York, NY, USA, 2005), ACM, pp. 1701–1704.

[110] Nooteboom, B. Learning to trust. In *Multidisciplinary Economics*, P. Gijsel and H. Schenk, Eds. Springer US, 2005, pp. 65–81.

[111] Patterson, O. Liberty against the democratic state: On the historical and contemporary sources of american distrust. In *Democracy and trust*, M. Warren, Ed. Cambridge University Press, 1999.

[112] Perrig, A., and Song, D. Hash visualization: A new technique to improve real world security. In *Proceedings of the International Workshop on Cryptographic Techniques and E-commerce* (1999).

[113] Pettit, P. The cunning of trust. *Philosophy and Public Affairs 24*, 3 (1995), 202–225.

[114] Phelps, E. A. Emotion and cognition: Insights from studies of the human amygdala. *Annual Review of Psychology 57*, 1 (2006), 27–53.

[115] Pillutla, M. M., Malhotra, D., and Murnighan, J. K. Attributions of trust and the calculus of reciprocity. *Journal of Experimental Social Psychology 39*, 5 (2003), 448 – 455.

[116] Putnam, R. *Making Democracy Work: Civic Traditions in Modern Italy*. Princeton University Press, 1993.

[117] Riegelsberger, J., and Angela Sasse, M. Trustbuilders and trustbusters. In *Towards the E-Society*, B. Schmid, K. Stanoevska-Slabeva, and V. Tschammer, Eds., vol. 74 of *IFIP International Federation for Information Processing*. Springer Boston, 2001, pp. 17–30.

[118] Riegelsberger, J., and Sasse, M. A. Face it - photos don't make a web site trustworthy. In *CHI '02: CHI '02 extended abstracts on Human factors in computing systems* (New York, NY, USA, 2002), ACM, pp. 742–743.

[119] Riegelsberger, J., Sasse, M. A., and McCarthy, J. D. The mechanics of trust: A framework for research and design. *International Journal of Human-Computer Studies 62*, 3 (2005), 381 – 422.

[120] Riegelsberger, J., Sasse, M. A., and McCarthy, J. D. The mechanics of trust: a framework for research and design. *Int. J. Hum.-Comput. Stud. 62*, 3 (2005), 381–422.

[121] Riekki, J., Salminen, T., and Alakarppa, I. Requesting pervasive services by touching rfid tags. *IEEE Pervasive Computing 5*, 1 (2006), 40.

[122] Rodden, T., Rogers, Y., Halloran, J., and Taylor, I. Designing novel interactional workspaces to support face to face consultations. In *CHI '03: Proceedings of the SIGCHI conference on Human factors in computing systems* (New York, NY, USA, 2003), ACM, pp. 57–64.

[123] Rogers, Y., and Price, S. Using ubiquitous computing to extend and enhance learning experiences. In *Ubiquitous Computing in Education: Invisible Technology, Visible Impact.*, M. V. t'Hooft and K. Swan, Eds. Sage Publications, 2006.

[124] Rotter, J. B. A new scale for the measurement of interpersonal trust. *Journal of Personality 35*, 4 (1967), 651–665.

[125] Rousseau, D. M., Sitkin, S. B., Burt, R. S., and Camerer, C. Not so different after all: a cross-descipline view of trust. *Academy of Management Review 23*, 3 (July 1998), 393–404.

[126] Rutter, J. Sociology of trust towards a sociology of 'e-trust'. In *International Journal of New Product Development Innovation Management* (2001), pp. 371–385.

[127] Sabel, C. Studied trust: building new forms of cooperation in a volatile economy. *Human Relations 46*, 9 (1996), 1133–1170.

[128] Sasse, M. A., Brostoff, S., and Weirich, D. Transforming the 'weakest link' — a human/computer interaction approach to usable and effective security. *BT Technology Journal 19*, 3 (2001), 122–131.

[129] Satyanarayanan, M. Pervasive computing: Vision and challenges, August 2001.

[130] Schilit, B., Adams, N., and Want, R. Context-aware computing applications. In *In Proceedings of the Workshop on Mobile Computing Systems and Applications* (1994), IEEE Computer Society, pp. 85–90.

[131] Schultz, C. D. A trust framework model for situational contexts. In *PST '06: Proceedings of the 2006 International Conference on Privacy, Security and Trust* (New York, NY, USA, 2006), ACM, pp. 1–7.

[132] Shapiro, D. L., Sheppard, B. H., and Cheraskin, L. Business on a handshake. *Negotiation Journal 8*, 4 (1992), 365–377.

[133] Shelat, B., and Egger, F. N. What makes people trust online gambling sites? In *CHI '02: CHI '02 extended abstracts on Human factors in computing systems* (New York, NY, USA, 2002), ACM, pp. 852–853.

[134] Sheng, S., Magnien, B., Kumaraguru, P., Acquisti, A., Cranor, L. F., Hong, J., and Nunge, E. Anti-phishing phil: the design and evaluation of a game that teaches people not to fall for phish. In *SOUPS '07: Proceedings of the 3rd symposium on Usable privacy and security* (New York, NY, USA, 2007), ACM, pp. 88–99.

[135] Sheppard, B. H., and Sherman, D. M. The grammars of trust: A model and general implications. *The Academy of Management Review 23*, 3 (July 1998), 422–437.

[136] Shneiderman, B. *Designing the User Interface: Strategies for Effective Human-Computer Interaction*. Addison-Wesley Longman Publishing Co., Inc., Boston, MA, USA, 1997.

[137] Simmel, G., and Wolff, K. H. *The Sociology of Georg Simmel*. Free Press, August 1964.

[138] Sitkin, S. B., and Roth, N. L. Explaining the Limited Effectiveness of Legalistic "Remedies" for Trust/Distrust. *ORGANIZATION SCIENCE 4*, 3 (1993), 367–392.

[139] Smetters, D. B., Balfanz, D., Smetters, D. K., Stewart, P., and Wong, H. C. Talking to strangers: Authentication in ad-hoc wireless networks.

[140] Spasojevic, M., and Kindberg, T. Evaluating the cooltown user experience. *Workshop on Evaluation Methodologies for Ubiquitous Computing at UbiComp'01* (2001).

[141] Stajano, F., and Anderson, R. The resurrecting duckling: Security issues for ad-hoc wireless networks. Springer-Verlag, pp. 172–194.

[142] Steinbrück, U., Schaumburg, H., Duda, S., and Krüger, T. A picture says more than a thousand words: photographs as trust builders in e-commerce websites. In *CHI '02: CHI '02 extended abstracts on Human factors in computing systems* (New York, NY, USA, 2002), ACM, pp. 748–749.

[143] Stolle, D. Bowling together, bowling alone: The development of generalized trust in voluntary associations. *Political Psychology 19*, 3 (1998), 497–525.

[144] Suchman, L. Plans and situated actions: Cambridge university press.

[145] Tan, Y.-H., and Thoen, W. A logical model of trust in electronic commerce.

[146] Uslaner, E. M. Producing and consuming trust. *Political Science Quarterly 115*, 4 (2000), 569–590.

[147] Uslaner, E. M. *The Moral Foundations of Trust*. Cambridge: Cambridge University Press, 2002.

[148] Varshavsky, A., Scannell, A., LaMarca, A., and De Lara, E. Amigo: proximity-based authentication of mobile devices. In *UbiComp'07: Proceedings of the 9th international conference on Ubiquitous computing* (Berlin, Heidelberg, 2007), Springer-Verlag, pp. 253–270.

[149] Viega, J., Kohno, T., and Potter, B. Trust (and mistrust) in secure applications. *Commun. ACM 44*, 2 (2001), 31–36.

[150] Weiser, M. Ubiquitous computing. *Computer 26* (1993), 71–72.

[151] Williamson, O. *The economic institutions of capitalism*. New York: Free Press, 1985.

[152] Winston, J., Strange, B., O'Doherty, J., and Dolan, R. Automatic and intentional brain responses during evaluation of trustworthiness of faces. *Nature Neuroscience 5*, 3 (2002), 277–283.

[153] Yousafzai, S. Y., Pallister, J. G., and Foxall, G. R. A proposed model of e-trust for electronic banking. *Technovation 23*, 11 (2003), 847 – 860.

[154] Zajonc, R. B. Feelings and thinking: Preferences need no inferences. *American Psychologist 35* (1980), 151–175.

[155] Zak, P. J., and Knack, S. Trust and growth.

[156] Zambonelli, F. Pervasive urban crowdsourcing: Visions and challenges. In *In Proceeding of PERCOM '11, IEEE Computer Society* (2011).

[157] Zheng, J., Veinott, E., Bos, N., Olson, J. S., and Olson, G. M. Trust without touch: jumpstarting long-distance trust with initial social activities. In *CHI '02: Proceedings of the SIGCHI conference on Human factors in computing systems* (New York, NY, USA, 2002), ACM, pp. 141–146.

[158] Zhu, F., Mutka, M. W., and Ni, L. M. Service discovery in pervasive computing environments. *IEEE Pervasive Computing 4* (2005), 81–90.