Evaluating usability of e-commerce sites by tracking eye movements

Thesis

Evaluating Usability of E-commerce Sites by Tracking Eye Movements

The Open University
Faculty of Maths and Computing
Department of Computing

Milton Keynes
Walton Hall
United Kingdom
MK7 6AA

The Open University

Evaluating Usability of E-commerce Sites by Tracking Eye Movements

MSc Applied Psychology & Human Factors, Cranfield University, 2000
BSc (Hons) Psychology & Philosophy, University of Wales, Swansea, 1999

Submitted in partial completion of the requirements for the degree of Doctor of Philosophy in Human-Computer Interaction

December, 2005

Supervisors: Dr. Shailey Minocha, Dr. Marian Petre, Dr. Andy Grayson

Acknowledgements

Back in 1996 I set off for a challenge to explore my own path. I left my country, Greece, to study in the UK. Swansea was my first stop (BSc), then came Cranfield (MSc), then came the Open University. I am grateful to the scholars who contributed to my knowledge and inspired me to go on.

My first experience using an eye tracking device was at Enterprise IDU (www.enterpriseidu.com). Thanks to Dr Rob Waller, David Lewis and Dr. Judy Delin for offering me my first job as a usability researcher and introducing me to the world of eye tracking. I also had the chance to work with Laura Cowen with whom I shared a in eye tracking.

The first person who suggested to me to do a PhD was Professor Pat Hall of the OU (Open University). I owe him gratitude for his faith in me and for giving me the idea to explore the path of research. I am greatly indebted to three key people, my supervisors: Dr. Shailey Minocha, Dr. Marian Petre and Dr. Andy Grayson. Each one has advised me in different ways. A very big thank you for their patience, encouragement and support. They have been continually enthusiastic, inspirational, generous and kind. Many thanks to other people within the OU for support and useful advice: Debbie Briggs, Dr. Trevor Collins, Maja Dunn, Dr. Alvaro Faria, Dr. Max Garagnani**,** Merisa Hooton, Prof. Darrel Ince, Jennifer Rook, Prof. Wendy Stainton-Rogers, Dr. Judith Segal, and Dr. Helen Sharp.

All the way along I had the chance to meet important researchers who have helped me in one way or another. I would like to start off by thanking my research buddy and personal friend Jonathan San Diego of the Institute of Educational Technology at the OU for initiating critical discussions and giving me a hard time when talking about my work. He has contributed towards the development of my argumentation skills! Many thanks to my fellow researchers at the OU: Dr. Jose Abdelnour Nocera, Ann Abraham, Dr. Yoseph Araya, Dr. Antonio Barrios, Dr. Andrea Capiluppi, Francis Chantree, Dr. Liisa Dawson, Albert Flores, Dilek Hacialioglu, Debra and Charles Haley, Anne Hewling, Israel Herraiz, Dr. Pejman Iravani, Iestyn Jowers, Dr. Dinar Kale, Angela Lozano-Rodriquez, Phebe Mann, Miquel Prats, Avik Sarkar, Joan Serras at the OU for putting-up with my grumpy face during my lowest moments. Special thanks to Geke Van Dijk and Chandra Harrison for proof reading the dissertation.

Thank you to those I met in HCI conferences who spent time with me discussing about my research: Dr. Sally Fincher, Prof. Thomas Green, Prof. Robert Jacob, Dr. Ewald Kaluscha, Dr. John McCarthy, Dr. Sine McDougall, Sunila Mondi, Dr. Harry Narayanan, Dr. Tony J. Renshaw, Dr. Jens Riegelsberger, Dr. Gordon Rugg, Dr. Dan Russell, Natalie Webb, and Prof. Daniela Zambarbieri.

I would like to also thank the chairs of the Doctoral Consortia at the Interact 2003 Conference (Prof. Mark Apperley, Prof. Alistair Sutcliffe, Dr. Fabio Paterno and Dr. Phil Gray) and at the BCS-HCI 2003 Conference (Prof. Ann Blandford and Dr. Lachlan MacKinnon). In addition I am grateful to the anonymous reviewers, who provided constructive comments on the papers describing this work.

Thank you to Dr. Jakob Nielsen who showed interest in my work and invited me to attend his User Experience 2005 Conference.

Thank you also to Prof. Jeff Johnson from the Faculty of Technology, Design and Innovation of the OU for inviting me to present my work in a departmental seminar session and for providing feedback.

During my PhD studies I had the chance to co-supervise a BSc student, Selamawit Mebrahtu from Vesalius College, University of Brussels, who wrote her thesis on eye tracking. Thank you Selam for giving me this chance, I thoroughly enjoyed it.

I would like to thank all the participants who voluntarily participated in my studies and showed enthusiasm for my research topic. This research was funded in part from the Engineering and Physical Sciences. Additional subsistence expenses were kindly funded by the Research School and the Department of Computing of the OU.

Thank you to the Faculty of Education and Language Studies/Childhood Development and Learning of the OU for allowing me to use their eye tracking devices and their lab. I also received support and help from Eileen Mansfield and Chris Golding.

Despite the distance and the years I have been away there are people from Greece who still support me, remember me, believe in me and constantly nag me to go back home! Thank you to my good friends: Vaso Anagnostopoulou, Kostas Bouyioukolaptis, Tasia Gerostathopoulou, Alexandros Isaias, Rena Matalliotaki, Marianna Maurona, Natasha Nakou, Effie Orphanidou, Stauros Pothos, Georgios Tsolis, Stella Ventouri.

To my family I owe the most. There is no way I can repay the support they have offered me during all the years I have been chasing my dream. I thank them for bringing me up in a way that I am proud of. Thank you to my mother for being my number one fan, her faith in me gives me courage to reach for the sky. Thank you to my father who taught me not to be afraid to speak-up and defend myself. Thank you to my brother for reminding me that research is not the only important thing in life. Last but not least, special thanks to Des who has stood next to me throughout my studies and helped me immensely. I am grateful to all who have contributed in one way or another during the course of this PhD.

Related Publications

Tzanidou, E., Petre, M., Minocha, S., and Grayson, A. (**2005**)."Combining Eye Tracking and Conventional Techniques for Indications of User-Adaptability". In Costabile, M. F. and Paterno, F., (Eds). Proceedings of Interact 2005, IFIP TC13 International Conference on Human-Computer Interaction, LNCS, Springer – Verlag, Italy, pp. 753-766.

Tzanidou, E., Minocha, S., Petre, M., and Grayson, A. (**2005**). "Revisiting Web Design Guidelines by Exploring Users' Expectations, Preferences and Visual Search Behaviour". In McEwan, T., Gulliksen, J., and Benyon, D. (Eds). Proceedings of 19th British HCI Group Annual Conference, LNCS, Springer-Verlag, Edinburgh, pp. 421-438.

Tzanidou, E., Minocha, S., Petre, M., (**2005**). "Applying Eye Tracking for Usability Evaluations of e-commerce Sites ". Presented at the workshop on 'Commercial Uses of Eye tracking' held at the 19th British HCI Group Annual Conference, Napier University, Edinburgh; paper available at URL - http://www.amber-light.co.uk/HCI2005/Eyetracking/index.htm

Tzanidou, E., Minocha, S., Petre, M., Grayson, A., (**2004**). "An Examination of e-commerce Homepage Design Guidelines by Measuring Eye Movements". In Dunican and Green (Eds). Proceedings of the Psychology of Programming Interest Group (www.ppig.org), 16th Annual Workshop, Institute of Technology, Ireland, pp. 133-139.

Tzanidou, E. (**2003**). "Evaluation of e-commerce Sites by Tracking Eye Movements", Poster Presentation at the 6th HCI Educators Workshop 2003: Effective Teaching and Training in HCI. In Cairncross, S., Varey, A., and McEwan, T. Proceedings of the 6th HCI Educators Workshop, Napier University, Edinburgh, pp. 54-55.

Tzanidou, E. (**2003**). "Eye tracking as a complementary evaluation technique for e-commerce sites". Proceedings of the workshop on 'Exploring The Total Customer Experience (TCE): Usability Evaluations of (B2B) e-commerce Environments' held at INTERACT 2003, Zurich, Switzerland; paper available at URL - http://computing.open.ac.uk/interact2003

Tzanidou, E. (**2003**). "Improving Usability of e-commerce sites by Tracking Eye Movements". In Gray, P., Johnson, H., and O'Neil, E. (Eds). Proceedings Volume 2, British HCI Group, 17th British HCI Conference. Presented at the Doctoral Consortium, HCI 2003 (www.bcs-hci.org.uk/hci2003), Bath, UK, pp. 143-144.

Tzanidou, E. (**2003**). "Improving Usability of e-commerce sites by Tracking Eye Movements". In Rauterberg, M., Mennozzi, M., and Wesson, J., (Eds). Proceedings of Interact 2003, IFIP TC13 International Conference on Human-Computer Interaction, IOS Pess. Presented at the Doctoral Consortium, Interact 2003 (www.interact2003.org), Switzerland, pp. 676-678.

Abstract

The majority of existing e-commerce design guidelines has been derived by conducting heuristic evaluations, without reporting the involvement of the users themselves. This research provides clarification on a number of existing web design guidelines for e-commerce sites based on empirical studies with users. Four studies were conducted and each study focused on a specific set of design guidelines as found in the literature. A combined qualitative and quantitative approach has been used, including a state-of-the-art technique, eye tracking. The eye movement data were complemented by user-profile data elicited through background questionnaires and user-perception data as captured through semi-structured interviews. The first study investigated users' initial impressions of homepages of e-commerce sites. The second study examined users' adaptability to persistent or varied placement of design elements. The third and fourth studies explored the effect of the presentation format of e-commerce web pages: the first in terms of the proportion of images, and the second in terms of how key icons related to an e-commerce transaction were presented. On the whole, the results of the studies corroborated existing design guidelines, but they also identified potential refinements. The thesis contributes both methodologically and empirically to Human-Computer Interaction. The combined methodological approach enables insight into the user experience that spans behavioural aspects such as visual search behaviour and visual search performance data, and subjective aspects such as user expectations and preferences. The empirical outcomes amplify the design guidelines from a user's perspective.

Preface

"Do you think, fellow citizens, that any man would ever have been willing to train for the pancratium or any other of the harder contests in the Olympic Games if the crown were given, not to the best man, but to the man who had successfully intrigued for it? No man would ever have been willing. But as it is, because the reward is rare and because of the competition and the honour, and the undying fame that victory brings, men are willing to risk their bodies, and at the cost of the most severe discipline to carry the struggle to the end."

Aeschines (390-314 BC)

Contents

List of Figures

List of Tables

Chapter 1 Introduction

It is just that we should be grateful, not only to those with whose views we may agree, but also to those who have expressed more superficial views, for those also contributed something, by developing before us the powers of thought.
Aristotle (384-322 BC)

1.1 The thesis

The majority of existing e-commerce design guidelines has been derived from heuristic evaluations, without direct reference to the users themselves. By conducting studies of user behaviour, performance and perception regarding e-commerce websites, we can test and refine design guidelines that have been derived from such expert inspection. There has been a growing interest in tracking users' eye movements as a means of evaluating user interfaces (Dix *et al*. 2004). This research proposes the combined use of eye tracking and other techniques such as background questionnaires and pre- and post-session interviews as an effective means of conducting usability evaluations of e-commerce websites.

1.2 Background and motivation

Statistics show that 67 per cent of transactions on the Web are never completed (Cohen, 1999). Only 36 per cent of customers are satisfied by electronic transactions, and this bad experience tends to drive customers to other channels such as call centres, catalogues or 'bricks and mortar' stores (Chatham, 2002). Organisations need to ensure that they distinguish themselves from their

competitors, and they also need to ensure that they design usable sites in order to retain customers and complete sales. The *Dot Com Survival Guide* (Hurst and Terry, 2000) discusses the failure of e-commerce sites and the amount of money lost by e-businesses. Although more organisations are becoming aware of the importance of providing a positive customer experience, many are actually slow to put their customer-centred strategies into practice.

It often happens that, due to restrictions of time and resources, organisations do not involve users or prospective customers, nor conduct usability evaluation prior to launching their e-commerce sites. Most organisations just ensure that their designers create sites according to well-known guidelines, assuming their sites will therefore be usable and successful. However, Preece *et al*. (2002), suggest that designers should not presume that applying design guidelines will guarantee optimal usability.

Although the HCI literature provides a variety of general web design guidelines (IBM, 2004; Nielsen, 2000) which are applied widely by practitioners – including application for e-commerce sites (Nielsen *et al*. 2001, Van Dyune *et al*. 2003) – web designers often experience difficulty applying them, as they can either be too general to interpret for a specific context, or too specific for more than one domain of e-commerce sites (Beier and Vaughan, 2003). While many design guidelines can apply to web design in general, there are particular pages of a site, such as homepages, that require additional attention. For example, Nielsen's *Alertbox* column (Nielsen, 2003) suggests that the ten most violated homepage design guidelines include the misuse of a liquid layout and graphics to show content. Although guidelines like these are straightforward to implement in a specific

context, they are often not supported with empirical evidence of user behaviour. The majority of existing web design guidelines has been derived by expert heuristic evaluations, without explicitly reported direct involvement of users (Ivory *et al.* 2001).

The motivation of this research is to test and elaborate existing e-commerce design guidelines by involving users in eye tracking studies. Each of the eye-tracking studies in this research focuses on a specific set of web design guidelines.

1.3 Research questions

The research objectives are to:

- provide empirical tests of existing web design guidelines for e-commerce sites.
- demonstrate the usefulness of eye tracking when combined with other **conventional techniques** to explore detailed user behaviour.

The main research question is:

What factors influence the users' visual search behaviour and performance when they look for specific information on an e-commerce site? That is, where do users look, and what do they see and interpret – and what aspects of a website influence that activity?

Figure 1-1 illustrates the research questions.

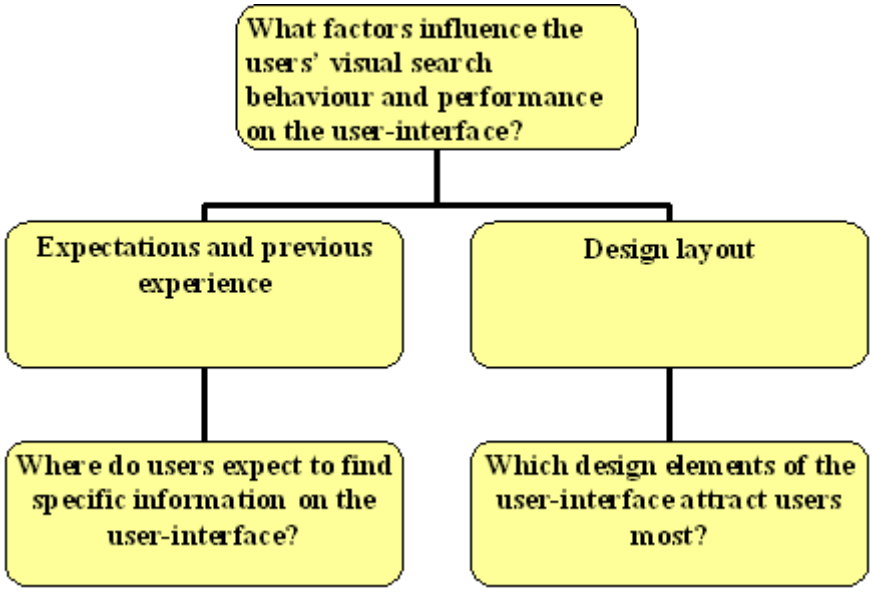


Figure 1-1 - Overview of research questions

The primary focus is on two main factors that influence the user's visual search:

- Expectations and previous experience
- Design layout and site attractors

Each of the studies conducted in this research has addressed a specific set of guidelines referring to factors that might influence visual search behaviour and performance. Table 1-1 illustrates the objectives of each Study.

Study #	Study name	Study's Objectives
1	Homepage study	• To explore the relationship between users' scanning behaviour and their ability to identify the brand and services on a home page. • To capture the relationship between users' previous experiences and expectations, and their scanning behaviour.
2	User adaptability study	• To examine how quickly users adapt to an unfamiliar design layout and, in particular, how quickly users adjust their expectations of where to look for a given target link during repeated exposures to a new design layout.
3	Text/image based study	• To explore how visual search behaviour is influenced by the presence or absence of text on homepages of e-travel sites. • To compare visual search behaviour across different design presentation formats of e-travel sites.
4	Shopping cart study	• To explore the effect of the presentation format (icon, such as the shopping cart, or written text, such as the words 'shopping cart') of the key steps in a transaction on an e-commerce site.

Table 1-1 - Summary of the eye-tracking studies and their objectives

1.4 Thesis structure

This thesis focuses on two aspects of HCI: a) the use of eye tracking as a usability evaluation technique and b) the assessment of existing web design guidelines. An outline of the thesis structure is presented below.

Chapter 1 introduces the background and motivation of this research. It goes on to state the research questions and contributions.

Chapter 2 initiates the discussion of web design guidelines. It discusses the impact of design guidelines for e-commerce websites. Next it reviews the evaluation techniques that are currently available to develop design guidelines. This review identifies several limitations of existing techniques. The remainder of the chapter sets the stage for using eye tracking as a technique.

Chapter 3 begins with a history of eye tracking and moves onto the use of eye tracking in HCI. It critically reviews up to date eye tracking studies, and provides the theoretical basis for the empirical research presented in this thesis.

Chapter 4 presents the methodology of the thesis. It presents the conceptual and terminological basis for the studies conducted in this research. Moreover, it includes the limitations of the chosen methodology and how this matter has been dealt with it in this thesis.

Chapter 5 presents Study 1, the *Homepage Study*. The aim of the Study was to capture the relationship between users' previous experiences and expectations in relation to their ability to identify the brand and services on a homepage. The homepage Study has been published and presented in the PPIG 2004 (Tzanidou *et al.* 2004) and the BCS-HCI2005 conferences (Tzanidou *et al.* 2005a).

Chapter 6 presents Study 2, the *User-adaptability Study*, which examines how quickly users adapt to an unfamiliar design layout and, in particular, how quickly the users adjust their expectations of where to look for a given target link during repeated exposures to a new design layout. The user-adaptability Study has been published and presented in the Interact 2005 conference (Tzanidou *et al.* 2005b).

Chapter 7 presents Study 3, the *Text/image Based Study*, which explores how visual search behaviour is influenced by the presence or absence of text on homepages of e-travel sites, and in particular the effect of the number of images appearing on e-travel homepages.

Chapter 8 presents Study 4, the *Shopping Cart Study*, which explored the effect of the presentation format (icon or a textual link) of the key steps in a transaction on an e-commerce site, such as the shopping cart.

Chapter 9 summarises the outcomes of the empirical studies and discusses their significance. It suggests recommendations for researchers and designers. The chapter closes with suggestions for future work.

The Glossary contains definitions of key terms as used in the dissertation.

The Appendices contain:

- Appendix A with *Evaluation Material* such as postscripts, questionnaires and consent forms, and
- Appendix B *with Data Consolidation and Results* tables.

Chapter 2 Web design guidelines

Then anyone who leaves behind him a written manual, and likewise anyone who receives it, in the belief that such writing will be clear and certain, must be exceedingly simple-minded.
Plato (427-347 BC)

2.1 Usability and design guidelines for e-commerce

Usability is said to be one of the most important determinants for the success of an e-commerce website (Nielsen *et al*. 2001). The failure of many e-commerce sites is attributed to lack of usability; for example, Herschlag (1998) suggests that customers do not shop online because they simply cannot find the item for which they are looking. The usability industry has responded by underlining the importance of designing clear websites, as reflected in general design guidelines. The HCI literature is rich with reports suggesting lists of web design guidelines for e-commerce websites (examples include: Nielsen *et al*. 2001, Spool 1999, Van Duyne *et al*. 2003) which are easily accessible by companies, designers and researchers. All agree on the need to apply the guidelines in order to achieve website usability.

An increasing number of e-commerce companies has realised the importance of designing usable websites. According to EPS IT Inc., by 2005 the number of internet users will reach 80 per cent of the US population. But the competition is huge, and it is so easy to switch from one site to another if people cannot find what they are looking for (Bhatti *et al*. 2000). E-commerce sites are the only sites that are re-designed regularly (Nielsen *et al*. 2001).

The assumption is that e-commerce usability should correspond to increased sales and increased customer loyalty. But how well do available design guidelines reflect the *customer's* perspective? To what extent will applying these guidelines lead to a more usable website that will meet the customers' needs (Limbourg and Vanderdonckt 2002)?

2.2 Customer-centred e-commerce website design

The users' perspective is of paramount importance to e-commerce companies, which depend on customers visiting their sites, making transactions and returning to the site in future. So a website designed well *for the user* is essential not just to facilitate transactions, but also in order to communicate effectively with customers. Communication covers a range of needs, from attracting their attention, through providing efficient service, to conveying the quality and reliability of the brand (Cao, 2002). A good site must engage users' beliefs and perceptions, in order to secure their purchases.

Hence, usability is not the only factor. Customers need to believe that the website with which they are interacting is trustworthy enough to deliver their purchase (Nielsen *et al*. 2001, Egger, 2003). Hence, customer behaviour, product perception, shopping experience and customer service must all be taken into consideration during design of an e-commerce website (Jarvenpaa and Todd, 1997).

"*When visitors give an e-commerce site high marks for content, ease of use, performance, trustworthiness and overall satisfaction, we call it a customer-centred website*" (Van Duyne, 2003). A customer–centred design is closely related to the traditional user-centred design first mentioned back in the 1980s for the development of usable computer-based systems.

Customer-centred design adds elements that go beyond the effective and efficient use of the website. The design approach focuses on marketing factors in relation to usability issues. When web design began, the initial focus was to design sites to impress users with flashy graphics and eye-catching images, without focusing on the customer's perspective. In recent years the focus has shifted to constructing powerful websites that provide real value and deliver a positive customer experience.

The adoption of the term 'customer' reflects the focus on issues such as brand value, trustworthiness and overall customer satisfaction. Beyer and Holtzblatt (1998) suggest the use of the term 'customer' to refer to anyone who uses or depends on a site. Customers can be administrators, business partners, managers, stakeholders and the target user population. The term customer is broader than 'user' and covers more individuals and the variety of their requirements. (Please note that 'user' and 'customer' are used inter-changeably in this dissertation.)

Van Duyne *et al.* (2003) suggest a list of the elements on which a customer-centred design should focus:

- Ease of use
- Content
- Performance
- Brand value
- Satisfaction

According to Van Duyne *et al*. (2003), overall, a customer-centred design strategy:

- Focuses on understanding people, their tasks, the technology available and the larger social and organisational context of where they live and work,
- Keeps the customer involved in the design process and,
- Elicits from visitors to a website consistently high marks for content, ease of use, performance, trustworthiness and overall satisfaction, it provides a positive experience for all customers, whether they are there to find information, to be part of a community, to purchase items or to be entertained.

2.3 How good are design guidelines?

It is assumed that design guidelines actually result in more usable sites, but most of the guidelines are not explicitly associated with evidence that following them will actually improve usability or accessability (Barrick *et al.* 2004, Vanderdonckt *et al*. 2003). Nielsen (1999) suggests that web design guidelines are well-established and have been corroborated repeatedly in many studies – however, those studies are not reported in association with the guidelines. In contrast, Spool *et al.* (2001) conducted empirical studies of guidelines which found that many guidelines which seemed obviously beneficial had no effect, while some actually reduced e-commerce sales.

There is evidently a need to provide designers with guidelines whose effectiveness and applicability has been demonstrated empirically. This has resulted

in attempts to develop web design guidelines based on user testing. *Usability.gov* provides an excellent list of 50 evidence-based guidelines, using evidence identified and compiled by the National Cancer Institute (NCI) in 2002. Each guideline provides a score indicating the strength of evidence that supports the guideline and examples of the guideline in practice. This helps determine the validity to attribute to each guideline, but while many guidelines claim to cite empirical evidence they simply cite other guidelines. Thus the connection between web design guidelines and the supporting research is often fragile (Barrick *et al.* 2004).

Despite this, the take-up of guidelines appears significant. When Nielsen *et al*. (2001) tested 20 e-commerce websites in 2000 for compliance with web design guidelines, only 45 per cent of the guidelines were followed. This number increased slightly in 2002 when they repeated the study with different websites. The average e-commerce website complied with 49 per cent of usability guidelines. If this level of improvement is sustained, Nielsen estimates that his ideal of 90 per cent compliance with the guidelines will be achieved by 2017 (Nielsen, 2002).

When testing the 'durability' of past web design guidelines, Nielsen (2005) found that 10 per cent of the guidelines would have to be retracted or reconsidered today. This suggests that some guidelines may not reflect enduring principles, or that they may be more culturally-sensitive than has been recognised. Moreover, Nielsen states that 20 per cent of guidelines are essentially irrelevant simply because they relate to rarely-used interface technologies. This durability test though is limited as it did not test the full list of original guidelines and some of the original websites tested had emerged in terms of design layout over the period of time since the first test.

The status of design guidelines is confusing: substantial compliance (49 per cent), but no guarantee of success through their application, and problematic durability (with 30 per cent of past guidelines judged inappropriate or irrelevant). If the guidelines are genuinely effective, why is compliance less than 49 per cent? On what are the guidelines based that might ensure effectiveness?

Guidelines come from various sources. Brinck *et al.* (2002) suggest that web design guidelines could be developed from broad concepts of design, user testing, practical intuition or from a psychological theory. Some guidelines are derived from ISO International Standards for example, IEC CDV TR 61997: *Guidelines for the user interfaces in multimedia equipment for general purpose use* (2000). These guidelines give general principles and detailed design guidance for media selection, and for mechanical, graphical and auditory user interfaces. Some have been developed based on years of experience of design and usability evaluation. For example, the guideline '*Do not rely on colour alone to communicate a message'* which requires the Federal Government Web sites to follow the Section 508 Federal (Web) Accessibility Standards (Chisholm *et al*. 1999). The majority of existing web design guidelines has been derived by expert heuristics as it is cheaper than conducting user-based studies (Ivory *et al*. 2001). Hence guidelines are not all equally well-founded, and they attract varying levels of confidence.

This combination of factors may well account for the current compliance levels; the lack of reported evidence, the uneven quality of the guidelines themselves, and the abundance of available web design guidelines from a variety of more- or less-plausible sources induces a credibility problem for practitioners. It is not just a matter of persuading companies to invest in usability, but also a matter of

demonstrating the validity of design guidelines and prioritising those which, when applied, will actually increase the website's success.

Jared Spool (1999) writes about the dangers of web design guidelines, based on his comparison of well-known guidelines to outcomes from his empirical studies. Spool *et al.* (2001) concluded that guidelines are very sensitive to the nature of the tasks and the subtle differences in the content of the site. Moreover, they suggest that a guideline's effectiveness will change even when the wording of the usability test's task is changed slightly. Spool *et al.* (2001) characteristically say: "*If we tell shoppers to buy a sweater, whether they are interested in buying one or not, we see different results from when we ask them to buy something they really need*".

2.4 Design guidelines are at different levels

Design guidelines, by their nature, are often difficult to develop and use (Tatzlaff and Shwartz, 1991). The variety of sources – and hence of perspectives – leads almost inevitably to different levels of design guidelines (Vanderdonckt *et al*. 2003):

- general usability guidelines that would apply to any design, for example: *make alert messages distinctive*

- domain-specific usability guidelines that apply to a specific type of site and, for example: *use a search box instead of a link to a search page*

- company-specific guidelines that apply to an individual site which depends on the exact products and services a company offers (Nielsen *et al*. 2001. p.20). For example: *allow the user to track stages of the delivery process*.

The most common problem is the ability to provide guidelines at the correct level of support: the guidelines are often criticised for being either too general to be applied in a particular case, or too specific to a particular domain, technology or task (Vanderdonckt *et al*. 2003).

For example, Fleming (1998) describes ten main principles of successful web navigation some of which are:

- include design elements that are easy to remember,
- keep consistency, provide feedback,
- have meaningful context, and support the users' goals,
- design differently for websites that are intended for community, learning, and information.

All these goals are well aligned with common sense, but they could be too general to apply in different domains of websites. Patel (2003) examined the effectiveness of web design conventions across different site types and concluded that there are differences in web design across different domains. For example, e-tail and e-travel websites stood out in that they were the fastest to download and had the least number of words. Therefore, with different site types, design guidelines may differ.

However, there are other web design guidelines that are more straightforward to implement and are broadly used by practitioners. For example, Nielsen (2003) claims that the most common mistakes of website design include using frames, long pages, non-standard link colours, and overly long download times.

2.5 Need to examine guidelines from a user's perspective

The issues surrounding design guidelines for e-commerce websites emphasise three key needs:

- Evidence: The diverse sources of guidelines and in particular the reliance on expert opinion to generate guidelines leaves a credibility gap. There is a need to link guidelines explicitly to evidence of effectiveness.

- Applicability: The different levels and perspectives reflected in the guidelines can make it difficult to identify pertinent guidelines and interpret them into practice.

- User perspective: Guidelines must genuinely reflect the user/customer needs, which extend beyond usability to address issues such as trustworthiness, loyalty, and satisfaction.

Overall, there is a need to apply user-centred techniques to re-examine e-commerce web design guidelines in order to fill the evidence gap and refine the guidelines for practical application.

2.6 Conventional usability evaluation techniques

This section reviews available data collection techniques with a view to determining the most effective method for revisiting existing web design guidelines. The list of techniques reviewed is representative, though not complete. Evaluation techniques can be categorised into two types, *User-based Observations* and *Expert Inspections* (Stone *et al.* 2005).

2.6.1 User-based observations

User-based observations involve the end user in providing feedback and identifying usability problems. Evaluators reason about the causes of the problems and make suggestions for redesign on the basis of the user information. The main focus of these techniques is to observe the users' behaviour with various aspects of the interface. Alternatively, observations measure users' task performance, typically using time to complete a task and errors made during it. Methods employed in user-based observations vary depending on the setting (for example, in the field or in the laboratory), the extent of the observer's involvement with the task (for example, two-way mirror or ethnographic approach), the formality of the observation process (for example, 'quick and dirty' or highly controlled), and the tools used to record the observations (for example, pen and paper or audio recording) (Preece *et al*. 2002). The following sub-sections will present a brief overview of some common user-based evaluation techniques: performance measurement, remote evaluation, interview, questionnaire, think-aloud protocol and retrospective verbal protocol.

Performance measurement

Performance measurement focuses on explicit user behaviour during the performance of a representative task. Typical performance measures include the time it takes a user to complete a given task (speed) and the number and type of errors made during a task (accuracy). The more quickly the user performs a task the more effective the user interface is assumed to be, and the fewer errors made during task completion the more efficient the user interface is. Speed and accuracy often trade-off. Ideally, performance observations should involve real customers,

completing realistic tasks in their natural work settings. For example, Minocha *et al*. (2003) conducted 'naturalistic' observations with self-motivated participants to explore their total customer experience. The artificial settings of a lab-based usability test may lead to data that do not accurately describe real-world behaviour (Reips, 2000). Performance observations are good for the detection of 'objective' phenomena but do not enable a better understanding of the cause of any performance differences. For example, efficient performance may not link to the user perception and experience.

Remote evaluation

In remote evaluation, the evaluator observes users who are not accessible, for example because they are in a different country. Remote evaluation is conducted via network or video conferencing facilities. Remote evaluation is also useful when the observer wants to simulate an environment where the users are in one location but the system is in another. There are several ways of collecting data in remote evaluations Hartson *et al.* (1996):

- Remote – control evaluation
- Video conferencing
- Remote questionnaire

Remote evaluation extends usability evaluation beyond the laboratory and potentially enables in situ observation of a broad range of users in their natural settings. It also makes the evaluator less evident to the user. However, remote evaluation requires additional software to observe participants from a distance which adds costs.

Interview

Interviews provide a rich source of data focused on the user's own account of the interaction. There are three main types of interviews: open-ended or unstructured (user-led), structured (following a strict script), semi-structured (following a script but allowing digressions). Interviews are usually conducted one-to-one, but group interviews can be used (Preece *et al*. 2002). The design of the interview depends on the evaluation goal and its specificity. Interviews are used to elicit information that direct observation might miss, because they probe the user's cognitive processes and opinions rather than just recording their behaviour. However, what users report depends on memory. Memories of events tend to be regularised, with details missing that might be of great importance (Schacter, 2001). Though interviews may not always be accurate, in many instances the users' preferences can only be revealed by asking the user to report them.

Questionnaire

Questionnaires can gather both quantitative and qualitative data (Oppenheim, 1992). Open or closed questions can be used depending on the evaluation goal. Closed questions provide a range of answers from which the user may choose, such as a specific scale. On the other hand, open questions allow the user to express him/herself to a greater extent without constraints. Well-phrased open questions do not imply a particular answer. Despite questionnaires being cheap to produce, they lack some of the flexibility that interviews offer. Questionnaires do not allow enough rapport to be established, which might result in unanswered or misunderstood questions (Oppenheim, 1992). Moreover, there is often a low return

rate on questionnaires, which introduces potential bias because of those who choose to respond.

Think aloud protocol

Think aloud aims to uncover the steps users take while interacting with the user interface. The user is asked to vocalise his/her thoughts, feelings, and opinions while performing a task as part of a user scenario. Think aloud provides greater understanding of how the user approaches the interface and what considerations the user keeps in mind when using the interface (Preece *et al*. 2002). However, it relies on the quality of the user's self-report and on what the user chooses to verbalise. It can only expose conscious, deliberate aspects of the user's experience. Verbalisation can interfere with time critical tasks and so is typically not compatible with time-based performance metrics.

Retrospective verbal protocol

Retrospective verbal protocols prompt participants to talk through how they completed a task while they watch a recording of what they were doing at the time. This technique aims to cue the participants' memories of events in a post-test walkthrough (Dix *et al*. 1998). However, there is a danger of post-hoc rationalisation; the participant might not be able accurately to remember the reasons for his or her actions but may still feel compelled to provide an account of them.

Conducting user-based observations allows the identification of problems arising from interaction with the user interface but can be time consuming in terms of both data collection and data analysis, depending on the numbers of users involved and the duration of the tasks. There are many concerns about bias, for example from the

influence of the particular task, the reliance on self-report, and the potential desire of the user to please the evaluator. Companies often prefer to conduct expert inspections for quicker results. The following section discusses benefits and disadvantages of expert inspections.

2.6.2 Expert inspections

An alternative to user-based techniques are inspection techniques which came into prominence in the early 1990s. Expert inspections, also known as expert evaluations, involve evaluators in examining or inspecting usability-related aspects of a user interface. Expert inspectors could be usability specialists, but could also be software development consultants with special expertise (for example, knowledge of a particular style of graphical interfaces), real end-users with specialist content or task knowledge, or other types of professionals (Nielsen and Mack, 1994).

In expert usability inspection, the evaluation of the user interface is based on the judgment of the inspector. There is no specific focus on the real end-user. Expert inspections are usually easier to conduct than user observations and require less time and resources to administer, but they can be more expensive. Moreover, expert inspections rely on small numbers of evaluators and therefore are sensitive to the individual bias. The following sub-sections will present a brief overview of some common expert evaluation techniques: heuristic evaluation, peer reviews, standard inspection, cognitive walkthrough, participatory evaluation and guideline reviews.

Heuristic Evaluation

Heuristic evaluation involves having a small set of evaluators examine the interface and judge its compliance with recognised usability principles called 'heuristics'. Jakob Nielsen's 1994 set of usability heuristics were developed with a focus on desktop applications. In 1997, Keith Instone (Instone, 1997) shared his thoughts on how these heuristics apply to what was a relatively new area: websites. Experts are guided by a set of heuristics that are developed for the evaluation of different systems. For example, Nielsen suggests that the following heuristics are more useful for evaluating commercial websites, and makes them memorable by introducing the acronym HOMERUN (Preece *et al*. 2002).

- High – quality content
- Often updated
- Minimal download time
- Ease of use
- Relevant to users' needs
- Unique to the online medium
- Net-centric corporate culture

Heuristic evaluations are less time-consuming to conduct than user-based observations. However, it is difficult to summarize the findings from multiple evaluators as they report problems differently and at different levels.

Cognitive Walkthrough

Cognitive walkthrough is a task-oriented technique in which the expert systematically explores the system functionalities from a task perspective. For example, the expert simulates step-by-step user behaviour for a given task. Several versions of cognitive walkthroughs exist including pluristic walkthrough where end-users, software developers or designers and usability experts go through the system discussing every single aspect of it (Barnard and Barnard, 1995). A study (Nielsen and Philips, 1993) found that heuristic evaluations were better predictors than cognitive walkthroughs and guideline based evaluations. This was compared to results from laboratory usability tests. However, none of these methods found more than 50 per cent of the problems discovered in laboratory testing. Moreover, cognitive walkthroughs often do not accurately predict problems that the real user encounters.

Peer reviews

During a peer review a colleague outside the design team, rather than a usability expert, reviews the user interface. A peer review is usually informal and involves asking a colleague to assess the interface at an early stage of the design. Peer reviews are quite often applied in companies for internal use as it is quick to conduct. However, there are issues involved as to the realism of the evaluation, such as for example, the choice of tasks and how representative they are of real users.

Guideline and Standards Reviews

Guideline reviews are inspections where an interface is checked for conformance with a comprehensive list of usability guidelines. It is considered as a combination of standard inspections and heuristic evaluation. The issue however is to know which

guidelines to follow. The application of the guideline depends on many factors including the type of application, the expertise of the end-users, and the environment in which the application will be used (Drury, 2004).

Standards inspections focus on specific standards. Inspectors go through a checklist that is provided in a formal manner. This requires prior knowledge, and therefore it is necessary to ask an expert who is familiar with the language used. Checklists have the limitation of being finite. Some issues not previously considered may not be on the list.

Participatory evaluation

Participatory evaluation involves the end user. The users are asked to inspect the user interface based on Nielsen's list of heuristics. Users are given the list as they do not have any prior HCI knowledge or expertise. User-centred design and consultation could be considered as a form of participatory evaluation but a lot of information about the users such as cultural differences needs to be gathered and taken account for.

2.7 Review of usability evaluation techniques

There have been several attempts to critically review and compare existing evaluation techniques for the evaluation of user interfaces. A summary of these is provided by Hartson *et al*. (2000) who conducted a review of 14 comparison studies of the effectiveness of usability evaluation studies. Seven studies specifically compared the heuristic evaluation technique with other usability evaluation techniques. Heuristic evaluations were reported as having a higher thoroughness rating in six out of these seven studies (85.7 per cent). Based on this high

percentage of thoroughness of problem identification, heuristic evaluation appears to reveal more problems than other usability evaluation techniques when compared head to head. Such a conclusion is often reported in the literature with only a few exceptions (Hartson *et al*. 2000). However, heuristic evaluations carry the limitation of not involving real users which may lead to the identification of a higher proportion of lower severity problems. The seven studies are discussed in more detail in the following paragraphs.

2.7.1 Heuristic evaluation vs. cognitive walkthrough

In his review Hartson *et al*. (2000) report that Dutt *et al.* (1994) focused on the evaluation of the effectiveness of heuristic evaluation and cognitive walkthrough for the identification of usability problems in a database system with complex interface functionality. The study found that, in terms of the number of usability problems, heuristic evaluation identified more than the cognitive walkthrough. However it is not established whether the problems identified are representative of the real users. The results suggest that the two techniques should be employed as complements to each other and at different stages of the design process.

2.7.2 Heuristic evaluation vs. user-based observations

Nielsen and Phillips (1993) compared heuristic evaluation and user-based observation techniques in predicting user performance. Their results suggested that heuristic evaluation is cheaper to perform and, if conducted by more than one usability specialist, it identifies more of the serious problems than any other evaluation method. But that mean values of at least three expert evaluation estimates were reasonably close to the values measured user-based observations.

2.7.3 Heuristic evaluation validated through performance measures

Borges *et al.* (1996) compiled guidelines for the design of web pages, from heuristic evaluation of existing websites, and he then evaluated the redesigned websites to assess the validity of the guidelines derived from heuristic evaluation. The heuristic evaluations, often of university and college homepages, revealed a variety of usability problems. A list of guidelines was derived to address the usability problems encountered. In order to evaluate the proposed guidelines, an experiment was conducted where versions of the redesigned web pages, based on the new list of guidelines, were evaluated. Typical users were asked to complete a set of tasks on both the original versions of the web pages and the latest versions while time for the task completion was measured. The results revealed that the average time to complete a large majority of the tasks on the redesigned homepages was significantly reduced validating their guidelines. The study suggests that web designers can improve the usability of home pages by applying guidelines from heuristic evaluations.

2.7.4 Heuristic evaluation vs. user testing

On the contrary a study by Fu *et al*. (1998) suggests that users are more effective than human factors experts in finding usability problems associated with the knowledge-based level of performance. Fu *et al*.(1998) evaluated web-based software used by two groups; human factors experts and users. Usability problems were classified into three categories associated with different levels of performance: skill-based, rule-based, knowledge-based. Human factors experts examined the interface and judged whether its design complied with recognised usability

heuristics. The user testing involved users interacting with a computer. The testing was conducted in a usability lab where user performance measurement, think aloud protocol collection and scenario based usability testing were applied.

A less conclusive study by Law and Hvannberg (2002) compared the effectiveness of heuristic evaluation and usability testing. Their results indicate that 19 out of 43 usability problems were identified by heuristic evaluation and 17 out of 43 were identified by usability testing. Their results suggest the complementary convergence of both techniques but do not indicate any conclusive explanation about their divergence.

2.7.5 Summary of comparisons

Review studies of the effectiveness of different evaluation techniques fail to meet a standard criterion to make valid comparisons. The majority of the comparison studies in the HCI literature on usability evaluation techniques effectiveness do not provide the descriptive statistics needed to perform a meta-analysis. Despite the fact that the majority of review studies conclude that heuristic evaluation identifies more usability problems compared to user-based observations, very few papers consider the effectiveness of the evaluation techniques in facilitating better re-design of the user interface. In addition, even though heuristic evaluations seem to be an easy and cheap way to evaluate the user interface, there is poor evidence of their ability to uncover real usability problems on complex user interfaces (Somervell and McCrickard, 2004). Moreover, other aspects of usability such as satisfaction are not taken into account when conducting heuristic evaluations. For e-commerce environments the user satisfaction can often be a determinant of the user's choice of a website.

Monk (1998) points out that comparative experimental studies are intended for designing small questions among same things, not grand cross questions such as deciding which usability evaluation technique to use and why one is better than the other. The only thing that usability evaluation techniques can be sure to have in common is that they produce lists with usability issues when applied to target usability problems. It is necessary to adjust the conditions for the application of usability problems detected. For example, one technique might be good for focusing an evaluation instance to narrow the scope of the inspection to just meet evaluation needs at a certain stage of product development that might represent a cost savings over another technique that would be more effective at a different stage (Hartson *et al*. 2000).

Therefore, each evaluation technique potentially identifies usability problems at a specific stage of the interface design, but when used in collaboration with other techniques, identification of usability problems will increase. Each technique is better than another depending on several factors, such as the evaluation goals (questions to be answered), the resources (money and time available), and the development stage of the user interface.

Chapter 3 Eye tracking in HCI

When the mind's eye rests on objects illuminated by truth and reality, it understands and comprehends them, and functions intelligently; but when it turns to the twilight world of change and decay, it can only form opinions, its vision is confused and its beliefs shifting, and it seems to lack intelligence.
Plato (427-347 BC)

3.1 History of eye tracking technology

Eye movements can provide a rich and detailed insight into a person's thoughts and intentions Rayner and Pollatsek, (1994). In recent years researchers have increasingly studied eye movements in order to explore cognition in psychology experiments, to understand user behaviour in user interfaces and to control computer-based devices through eye-based input (Jacob and Karn, 2003; Salvucci, 1999). However, the study of eye movements dates back to 1901 when Dodge and Cline (1901) developed the first eye tracking system which required participants not to move their heads and recorded only horizontal eye position onto a falling photographic plate. Dodge's method, as it became known, was the first to capture eye movements accurately and non-invasively. This basic technique was used during the 1970s (Taylor, 1971).

In the 1920s eye tracking systems were developed further to record two-dimensional eye movement records (Gilliland, 1921). The first bulk of eye tracking research focused on the relationship between imagery and eye movements (for example Clark, 1916; Goldthwait, 1933; Stoy, 1930). Later the focus moved into

the investigation of processes, habits, individual and cultural differences in reading (for example Walker, 1933; Stone, 1941).

The progress of digital technology and image processing in the 1970s brought a breakthrough in eye tracking systems. Until then eye movements were recorded in relation to the head. But in order to identify where a person is looking on a surface, the head needed to remain still. This restriction was reduced by simultaneously measuring eye movements and the **point of regard**, the point on a surface where the person was actually looking. This did not entirely eliminate the restriction on head movements, but, with assistance from **chin rests** and **high-back chairs**, slight movements of the head could be isolated from eye movements which could yield more accurate eye tracking (Richardson, 2004).

The first head-mounted eye tracker was invented in 1948 by Hartridge and Thompson. Until then eye tracking devices were designed in a very tight and restrictive way towards the participant. The design of eye tracking devices improved in 1960 when Macworth and Thomas (1962) developed devices, which were less instructive of the participants' head movements.

Therefore, much of the early eye tracking research (through the 1970s) focused on the technical improvement of the eye tracking devices and techniques in order to make eye tracking more accessible, more accurate and better-supported from data collection through analysis using computers. Since then, the technology has progressed to head-mounted eye trackers and the most sophisticated remote devices that do not require any physical contact between the equipment and the participant.

In recent years eye tracking technology has improved significantly and devices are less expensive (Dix *et al*. 2004). Modern eye tracking systems allow head movement, which is an advantage for usability testing, as the user can sit naturally in front of the computer. Such eye tracking devices can be either:

- mounted on the head of a user (for example SMI‘s Head-Mounted Eye Tracking Device, HED-II) or
- remote from the user (for example ASL 504 pan/tilt eye tracker system).

Further, eye tracking can now leave the laboratory, with the recent introduction of mobile devices (for example, ASL’s ‘Mobile Eye’ which incorporates the device into a spectacles frame).

In principle, head mounted devices allow a great deal of head movement, but they are nevertheless sensitive to large movements. They often obstruct a small portion of the user’s visual field, and the user is aware of the system recording their eye movements. Remote devices do not require any physical contact with the user. They usually consist of a miniature camera mounted next to or in front of the computer display. In some cases the eye camera is hidden on the screen (for example Tobii ET-17™) which enables a more natural set up of the evaluation session.

While engineering improved the technology, psychology began exploring how eye fixations reflected **cognitive processes**.

3.2 Linking eye tracking to cognition: Fitts

Back in 1950 Fitts *et al*. were the first to use eye tracking to relate eye movement to cognitive activity in aviation studies. They captured pilots' eye movements with cockpit-mounted mirrors and a movie camera for cockpit design testing (Jacob and Karn, 2003). Data analysis of the footage involved time-consuming frame-by-frame counting. Nevertheless, this pilot study contributed valid findings on the arrangement of flight instruments. The study measured the frequency and pattern of eye movements during different phases of a flight (Fitts *et al*. 1950). Such pioneering research was valuable in establishing conclusions about the relationship between eye movement data and cognitive activity. For example, Fitts *et al*. (1950) argue that frequency of fixations relates to the importance of the control, while duration of fixations relates to the ease of interpreting information.

Four years later, in 1954 Paul Fitts proposed a robust model of human psychomotor behaviour known as 'Fitts Law'. The model is based on time and distance. Fitts' law is a model of human movement, predicting the time required to move rapidly from a starting position to a final target area, as a function of the distance to the target and the size of the target (Caroll, 1998). It seems intuitive that movement time would be affected by the distance moved and the precision demanded by the size of the target to which one is moving. Fitts discovered that movement time was a logarithmic function of distance when target size was held constant, and that movement time was also a logarithmic function of target size when distance was held constant. Mathematically, Fitts' law is stated as follows:

$MT = a + b \log2 (2A/W)$ where:

MT = movement time

a,b = regression coefficients

A = distance of movement from start to target centre

W = width of the target

Fitts' law is used to model the act of *pointing*, both in the real world, for example, with a hand or finger and on computers, for example, with a mouse (Fitts, 1995). However, it was not intended to predict finely-controlled movement such as drawing or writing (Caroll, 1998).

One way of applying 'Fitts' Law' to interface design is to relate the size of buttons to the distance from the pointer. But this suggestion may act in opposition to other factors that make a larger difference on interface efficiency, such as organization and use of available screen real estate. So, there are tradeoffs when applying 'Fitts' Law' to interface designs.

3.3 Visual search behaviour

Supposing that a researcher is interested in the usability of a human-computer interface it is important to find ways of making sense of the eye movement data. As with any other topic in research, having a clear idea of what is to be investigated is a good start. But most importantly relating results of the investigation back to established theories helps understand eye movement data better.

Theories of visual search, as reported by Horowitz and Wolfe (1998), conclude that visual search relies on accumulating information over time about the identity of

design elements. Such theories enable designers to structure the user interface effectively and assist the user's visual search behaviour. **Post-cognitive modelling** research demonstrates that people use anticipatory location information to guide visual search, and that visual features sometimes guide the visual search (Hornof and Halverson, 2003). Despite extensive research into visual search behaviour in disciplines such as psychology, HCI researchers (for example Hornof and Halverson, 2003) have only recently underlined the importance of developing a unified understanding of users' visual search behaviour.

Visual information is thought to be processed in two ways (Card *et al.* 1999):

- *Pre-attentive* processing which is done rapidly through parallel processing and
- *Post-attentive* processing during which items (that is, words, objects) are looked at in a serial manner.

Visual search behaviour on websites is influenced by users' expectations about what is being looked for and where it might be located. Pirolli and Card (1995) discuss the design layout of the display as a bottom-up influence, and expectations as a top-down influence. Bottom–up processing refers to the design elements influencing the visual scene itself, such as presentation format, colour, and position, whereas top-down processing refers to the expectations the users develop such as their preferences and predictions when viewing a scene. Hence bottom-up influences may correspond largely to post-attentive processing, and top-down influences may correspond largely to pre-attentive processing. Interactions between top-down and bottom-up influences are identified as 'Information Scent' or 'Information Foraging' (Pirolli and Card, 1995). Unless the design elements such as

colour, menu items, graphs (bottom–up) are looked at, there is no 'scent', and therefore, there is no basis for selection.

3.4 Eye tracking and cognitive processes

The relationship between eye movements and **cognitive processes** has received substantial attention in recent years in several research domains. In psychology and cognitive sciences, eye movements have been studied to help elicit and analyse the cognitive activity in a variety of task domains. Just and Carpenter (1976) were the first in cognitive science to establish that gaze direction carries information about the focus of the user's attention. They suggested this could make a good modality for proactive interaction between human and computer. Other significant results have been uncovered in the domains of reading (for example, Just and Carpenter, 1975, 1980; Rayner, 1983; Schilling, Rayner and Chumbley, 1998) and word location (O'Regan *et al.,* 1999; Zola, 1984). For example, Just and Carpenter (1980) captured the eye-movements of 14 college students who read 15 short excerpts from *Time* and *Newsweek* magazines. The participants were asked only to read normally and to recall what they could of each paragraph after it was finished. The authors found that readers fixated an average of 67.8 per cent of the words, with content words being fixated 83 per cent of the time and function words 38 per cent of the time. This work asserted that the reader tries to interpret each word in a text as it is encountered rather than holding it in memory and assigning meaning later.

There are several reasons why eye movements have become a source of so much interest. The principal reason is almost certainly technical, as eye tracking equipment has improved significantly (as discussed above). But other reasons lie in

the development and refinement of theory. Salvucci (1999) explains that eye movement protocols represent actions at a fine temporal grain size, typically on the order of 10 msec. Eye movements yield important clues to human behaviour at this temporal level, such as what information people use in problem solving and why or how long it takes to process the information given, and when people forget and review previously encoded information. To put it simply, eye movement data can tell us "what, where and for how long" (Zambarbieri, 2005). Eye movements allow making inferences about **cognitive processes**. It can be argued that eye movements do not entirely reveal a person's **cognitive processes**, but their flexibility means they can be used as a supplement to other sources of data.

3.5 Eye tracking for usability evaluation

Conventional usability evaluation techniques such as user- based observations focus more on users' activities and performance rather than the understanding of users' **cognitive processes** (Golberg *et al*. 2002). Possible cognitive differences underlying performance differences are often unaddressed (Chapter 2). Therefore, aspects of task performance such as screen navigation and selection of menu items can be captured, but inferences about **cognitive processes** are more difficult to justify. Eye tracking has been applied in human-computer interaction studies as an indicator of usability. The benefit of applying eye movement data is that it could be used to support recommendations for how a user interface should be changed, rather than a broad assessment of the interface's usability (Goldberg and Winchansky, 2003). Further, eye movement data can fuel theories of interaction that provide a deeper understanding of usability in the broadest sense.

Findings from eye tracking studies that explore navigation, search and other interactions with online applications have influenced overall user interface architecture, design layout of screens, size of design elements and so forth. Several key studies are summarised here in order to illustrate the nature and scope of that influence.

Recommendations for navigation aids were provided by Pirolli (2001) who applied eye tracking to explore users' navigation of network browsers. They used number and duration of fixations as indicators of importance of design elements and difficulty of information extraction. For example the traversal rate (that is the number of movements from one link to another) for visual links was found to be nearly twice as fast as for the traditional text-based tree browser. Studies like the one from Pirolli (2001) add to the body of knowledge of the effectiveness of eye movement-based metrics. The metrics of number and duration of fixations are the most used metrics in eye tracking studies as discussed in a review by Jacob and Karn (2003).

When looking at the searching of multi-column lists for specific target items, eye tracking revealed different scanning styles (Crosby and Peterson, 1991). Differences were found between novice and expert users. For example, experts were able to find the target link with fewer saccades than would be required if they fixated on every target. The scanning styles were: *Comparative* (scanning between columns to compare items), *Down and Up* (starting from the top of a column and then moving to the column next to it and starting from the bottom upwards), *Scan from top* (always starting from the top of any column), and *Exhaustive* (scanning all areas and columns). Such study (Crosby and Peterson, 1991), demonstrates the potential use of eye tracking as an indicator of searching strategies on websites.

Specific design recommendations for a prototype web-based development tool were developed by investigating search efficiency on multiple screens Goldberg *et al*. (2002). Results of this study suggest that design recommendations should emphasise that important 'portlets' should initially be contained on the upper-left or top of screens to minimise search time. The study used a within-participants design that might have enabled a learning effect across interfaces. Despite the design recommendations that were developed matching existing design guidelines derived from other traditional evaluation techniques; eye tracking was demonstrated as a user-centred technique for the development of guidelines.

Faraday and Sutcliffe (1996) tracked eye movements of participants viewing a multimedia information presentation. Based on an experiment with six participants viewing an 18 second presentation, they developed guidelines for multimedia presentation, based on eye movement data. All of the participants in their study shifted their attention as an arrow pointed to various objects. Eye tracking was demonstrated as a useful technique for capturing detailed user-behaviour. Similarly to previous work, this study led to the development of guidelines.

Crowe and Narayanan (2000) compared interfaces based on users' eye movements and actions. They demonstrated the advantages of combining interaction logs and eye movement data to analyse users' interactions with interfaces. This work is one of the few studies that suggest the combination of eye movement data and data from other sources for a better understanding of the user behaviour and perception.

Narayanan and Schrimpsher (2000) conducted a study to:

- evaluate the usefulness and usability of Hypermedia Information Presentations Systems (HIPS),

- develop software for the analysis of eye movement data, and

- apply an interactive algorithm visualisation designed for computer science undergraduates in a pilot experiment.

An ISCAN ETL – 400 Eye tracker at 60Hz was used. The results indicate that eye tracking is effective in analysing human interaction with multimedia presentations. It revealed information that could not have been obtained by traditional evaluation techniques such as user observation.

Renshaw *et al*. (2003) focused on the relationship of visual designs and eye movement to understand participants' visual attention. The experimental design consisted of a comparison of two graphical formats, one designed in accordance with and one in contravention of, established design guidelines. To assess the usability of each format eye movements, success rates and user satisfaction were collected. The authors applied existing **eye movement based metrics** as applied in previous studies (Cowen *et al.* 2002; Goldberg and Kotval, 1998) plus a new metric they developed 'gaze orientation' which measured a series of consecutive fixations within an area of interest. The apparatus of the experiment was an ASL 504 pan/tilt eye tracker system; at 50 Hz. Eye movement data were found to be consistent with more traditional performance measures. The authors concluded that graphs constructed using design guidelines have a significant advantage over other designs in terms of time to complete task, accuracy and user satisfaction. This study

demonstrated the use of eye tracking when applied as an indicator of guideline effectiveness for graphs.

Previous research, as reported above, has demonstrated the use of eye tracking as an indicator of usability. Moreover, studies that examined the effectiveness of eye movement-based metrics provide valuable information of how to apply specific metrics to interpret eye tracking results. The development of design guidelines for 'portlets' and multimedia presentations based on eye movement data strengthen the potential of the use of eye tracking for the development of design guidelines in other domains such as websites. The following section reviews eye tracking studies of websites.

3.6 Eye tracking for web design

Only a small number of studies have been conducted on eye movement behaviour on web pages (Granka *et al*. 2004). This section discusses examples of eye tracking studies for web design.

Visual preferences of text and images have been explored by two studies: Ellis *et al*. 1998; Lewenstein *et al*. 2002. Ellis *et al.* (1998) conducted a pilot study of web page design to assess the value of eye tracking as a usability evaluation technique. The eye movement data revealed detailed information such as the extent to which graphics were viewed and whether they distracted the users. The results of this study indicated that users completed the task more quickly on text-based screens; overall text was preferred more than images. Despite the authors' strong suggestion that eye tracking has the power to provide useful feedback for web designers, they have not validated their data by comparing outcomes of their results to outcomes from other traditional evaluation techniques.

The Stanford Poynter Project study (Lewenstein *et al*. 2000) examined how users read on-line and off-line news. They found that text was viewed more than images by readers who read on-line news, whereas the opposite applied for readers who read off-line news. This study demonstrates the usefulness of eye tracking in detecting fine user behaviour such as visual search strategies.

Josephson and Holmes (2002) recorded eye movement data of users who repeatedly viewed three websites each belonging to a different domain. Although they used real websites for their data collection, their aim here was to test Norton and Stark's (1971) scan path theory by string editing. The theory proposed the use of serial encoding for the integration of sensory and motor elements which guide the saccades of the eye for each visual image. Josephson and Holmes (2002) suggest that some users tend to develop a habitually-preferred visual path, scan path, across different displays of user interfaces. This is a descriptive study in nature, with no tests of significance. Similarly, McCarthy *et al*. (2003) tested whether the placement of menus on web pages in unexpected positions affected search performance, suggesting that users learn to adapt to screen layout that does not follow the **norms**.

The studies discussed above explored navigational styles that users apply on both off-line and on-line channels but did not investigate the factors that influence navigation styles. However, recent studies (for example Outing and Ruel, 2004, Pan *et al*. 2004) aim to understand the factors that influence user's visual-search behaviour.

Pan *et al*. (2004) investigated some of these factors, such as individual differences, design characteristics of the web pages, the order in which web pages are viewed and different tasks that were given to the users to complete. Gender and

viewing order were found to be key determinants of visual search behaviour. Men applied different scan paths from women and the order in which the stimuli were presented influenced the scan paths as well.

The Stanford Poynter Project (Outing and Ruel, 2004) extended their previous work (Lewenstein *et al*. 2000) on how users read news websites. They applied a more controlled approach in their latter study. Some key points of their latest study (Outing and Ruel, 2004) suggested that users navigate more on the upper part of news websites rather than on the left or right of the page. Text size was found to be influential in terms of encouraging focused viewing behaviour; smaller text drew more fixations while larger sizes promoted lighter scanning. The users fixated more on headlines with large text rather than headlines with small text.

Recent interest in the use of eye tracking in usability studies for the web have demonstrated the usefulness of eye tracking both as an indicator of usability but also as a detector of possible factors that might influence visual search behaviour. Moreover, by combining eye tracking with other techniques, richer data can be captured which enable a better understanding of the user's behaviour. The following section describes the proposed research, which aims to explore such influential factors in order to enrich existing web design guidelines.

3.7 Proposed research

Eye tracking has been classified as 'promising' over the last 50 years (Jacob and Karn, 2003) but recent technological advances have increased interest in its application. Nevertheless the HCI community must still learn more about users' deployment of visual attention and must learn more about designing product interfaces that fit human needs more closely (Jacob and Karn, 2003).

Goldberg and Wichansky (2003) underline the importance of more eye tracking research on factors that influence people's eye movements. For example (p.514): "Little is known about how the density of a display or visibility of icons influences eye tracking results and perceived usability".

The research reported in this thesis focuses exactly on such factors and how they influence visual search. Eye tracking has been applied as the primary research technique, but it is not used in isolation. Just as it has been suggested elsewhere (see Chapter 2) that different usability techniques can be combined for a fuller and more accurate evaluation, eye tracking too can benefit from combination with other techniques (Crowe and Narayanan, 2000). The research reported here combined eye tracking with interview and performance measures in order to triangulate results.

As discussed in Chapter 2 each evaluation technique has its strengths and weaknesses. Eye movement data represent detailed trails of visual search behaviour that can be analysed both qualitatively and qualitatively enabling powerful interpretations of the results. Eye tracking addresses some of the limitations of other techniques (improving precision and directness of capture) and when combined with other techniques it has the potential of producing a richer set of data. Studies that do not use eye tracking typically ask participants to report the difficulties they face during the user interface interaction, with the attendant issues of self-report. Eye tracking, however, is an objective way of addressing some of these issues, because it does not rely on memory or report, but records activity directly. On the other hand, eye tracking data must be interpreted. Changes in eye movements may reflect cognitive processes – or they may not. Differences in

visual search behaviour may signify difficulties in user-interface interaction, but the link must still be established.

Jacob and Karn (2003) explain that eye movement data can be analysed in three ways:

- *top-down,* when aspects of cognitive theories are tested,
- *top-down,* when a design hypothesis is tested, or
- *bottom-up,* when inferences are based on patterns across eye movement data.

The reported research has applied both top-down (hypothesis testing) and bottom-up (inductive) approaches, depending on the specific research question of each study (reported in Chapters 5 through 8). Guidelines are used as hypotheses, but also, the behaviour is examined in detail in order to understand *why* the guideline succeeds or fails.

As discussed previously (Chapter 1) the research objective is to validate guidelines based in user behaviour and experience. No one technique covers all aspects. Therefore, eye tracking is needed to enable direct access to user behaviour at fine granularity. Performance data is needed to relate eye movement behaviour to success in typical tasks. And interview and questionnaire data are needed to relate eye movement behaviour to user expectation and experience. Only such a combination of techniques, anchored in directly-captured behavioural data, provides sufficient information to test guidelines in detail.

Chapter 4 Methodology

A likely impossibility is always preferable to an unconvincing possibility.
Aristotle (384-322 BC)

4.1 Introduction

The previous chapters introduce the rationale of this research, outline available usability evaluation techniques, and focus on the primary evaluation techniques applied here. This chapter presents the methodology of all four studies. Although each of the reported studies is unique in terms of its design and objectives, all four share a common methodology. The set of techniques chosen for data collection are common across the four studies. The design sequence follows a similar pattern for all studies: a set of existing design guidelines is identified and then tested by conducting an eye tracking study (see Figure 4-2, for overview of research).

4.2 The techniques

This thesis takes the view that quantitative and qualitative approaches complement each other. Although there has been much debate about the dichotomy between quantitative and qualitative paradigms (Hammersley, 1999), research can be a pragmatic combination of quantitative and qualitative methods (Savenye and Robinson, 1996), in which the findings from each technique shed light on the findings from the other. This sort of 'triangulation' is used deliberately in human studies to explore the relationship between experience and behaviour (Mackay, and Fayard, (1997). Crowe and Narayanan (2000) suggest that the combination of eye tracking and other evaluation techniques gives a richer set of data.

The primary evaluation technique used in this research as an indicator of usability is eye tracking. Eye tracking data are complemented by data from background questionnaires and pre- and post-session interviews (Figure 4-1).

Background Questionnaire

↓

Pre-session Interview

↓

Eye Tracking Session

↓

Post-session Interview

Figure 4-1 - Collection of usability evaluation techniques as applied across all studies of this research

All four studies shared a common procedure for data collection. The participant was first asked to complete a background questionnaire, then to reply to the pre-session interview questions, then to complete tasks on web pages while the participant's eye movements were recorded, and finally to reply to the post-session interview questions. There were only minor differences in the procedure for data collection among the four studies. Study 1, the Homepage Study, did not have a pre-session interview. Study 1 prompted participants to give written examples of previously-used websites in the background questionnaire, whereas in the other three studies participants were asked to give examples during the pre-session interview. Figure 4-2 presents an overview of the research.

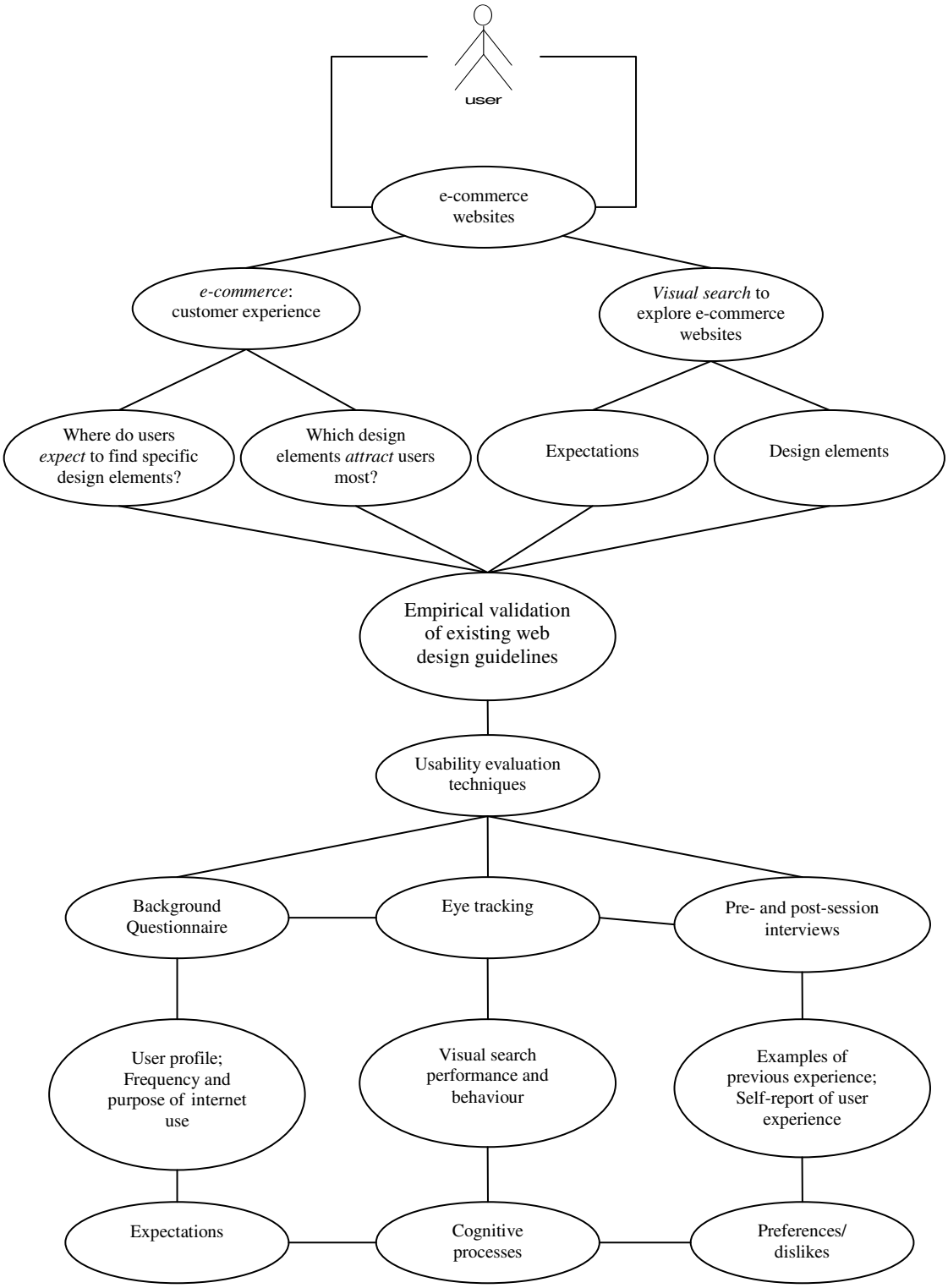


Figure 4-2 - Overview of research

4.2.1 Background questionnaire

The background questionnaire elicited:

- the user profile (that is: age, gender), and
- previous internet experience (such as frequency and purpose of internet use and, in study 1 only, examples of familiar web sites).

An example of the background questionnaire is found in Appendix A.

4.2.2 Pre-session interview

The pre-session semi-structured interview (studies 2, 3, 4) used open questions about users' expectations, preferences and dislikes regarding web pages. Typical questions asked participants where they had seen specific design elements placed on other websites and where they would like to find them placed on the user-interface. Interviews were audio recorded and later transcribed verbatim. Examples of pre-session interview protocols for studies 2-4 are found in Appendix A.

4.2.3 Eye tracking session

There is a need to ensure a common understanding of the use of terms and definitions related to eye tracking in the context in which they are used. For example, there is no standard operational definition for a single fixation, (where the eye stops briefly over a point), and there are different psychological theories about the relationship between eye movements and **cognitive processes** (Hansen, 2003). This section discusses the eye movement-based metrics that were applied across the four studies reported in this research. The definitions of the metrics are provided in

this section, but how each metric is applied to answer specific study questions is clarified in Chapters 5, 6, 7, and 8 for each study respectively.

Eye trackers provide a continuous stream of *x* and *y* coordinates, identifying the users' position of gaze at any specific moment. This data stream is then processed for interpretation. The eye movement model most commonly used is the fixation-saccade model. It divides eye movements into rapid movements (saccades) between points of relatively stable gaze (fixations) (Jacob and Karn, 2003; McCarthy, 2003).

The key eye tracking terminology was defined for these studies as:

Fixation: when the eye remains relatively stable. Published usability studies addressing visual search processing employ definitions of fixations with a minimum duration for a stable eye position ranging from 150-300 msec (Cowen *et al.* 2002). Depending on the type of eye tracking technology used and in particular the sampling rate of the data, definitions of fixation durations might vary across studies. Thus, it is advisable to first analyse raw eye-position data by defining how stable the eye must be and for how long it must be stable to qualify as a fixation. Fixation is used as an indication of information extraction (Hansen, 1990).

Saccade: ballistic eye movements, representing changes of eye position between fixations. A saccade is quicker than a fixation; it typically lasts between 50-150 msec. The duration of a saccade depends on its amplitude. The amplitude of a saccade is the angular distance that the eye needs to travel. A saccade is quite fast compared to any motion on the scene being observed, thus no information extraction takes place during a saccade. Saccades are used as an indication of scanning or shift of attention (Cowen *et al.* 2002).

Scan paths: the spatial arrangement of a sequence of fixations and saccades, that is, how the gaze moves and pauses over a period of time. Figure 4-3 represents an example of a user's scan path on a web page. The circles are fixations, the numbers inside the circles represent the duration and order of fixations and the lines between the fixations are saccades. Scan paths are used as an indication of the visual search sequence.

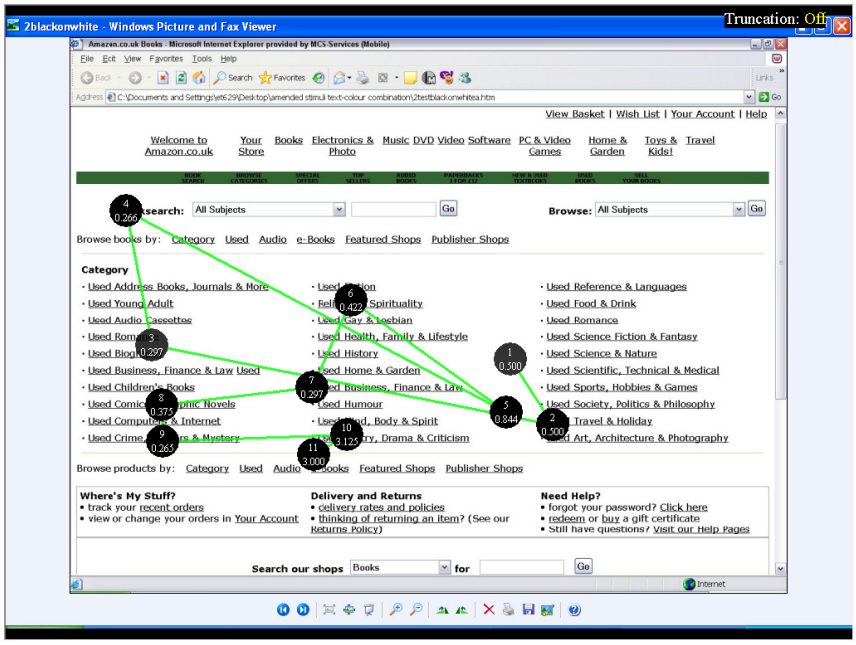


Figure 4-3 - Example of user's scan path

Look Zones or Areas of Interest: a selected display region whose boundaries are determined by design elements it contains. Hence *look zones* are determined 'semantically' to correspond to meaningful areas of the display. For example, a menu item, an image, or the banner on the user-interface might be defined as a look zone in order to determine whether, how often and for how long a user looks at that specific design element. Any fixations whose coordinates lie within the specified bounded region are used as an indication of attention to that design element.

Fixation, saccade, scan path and *look zones* are all phenomena which are commonly interpreted into straightforward count, duration and sequence metrics. Additional eye movement-based metrics were adopted – and in some cases adapted – for the four studies of the reported research are presented below:

Scan path efficiency: whether or not scan paths reach the bottom of the screen display. Rayner and Pollatsek (1994) found that well-organised grouping of component buttons in a screen display resulted in scan paths that covered smaller areas (that is, they resulted in more efficient scan paths). Hence, efficient (localised) scan paths indicate that the information is well organised and easy to find. In contrast, scan paths that reach the bottom of the screen display indicate that a comprehensive search was required and hence that the organisation is less efficient.

Initial gaze: where the participant looked less than a second (100 msec) prior to the web page appearing on the screen. It was suggested by Loftus and Mackworth (1978), who conducted eye tracking studies in reading that the eyes fixate initially on areas that are surprising, salient or important through experience. Hence location of initial fixation is significant. In this research, location of initial fixation has been separated into two metrics, of which initial gaze is the first (and entry point, below, is the second). It follows from Loftus and Mackworth that where participants look before any stimulus appears but when stimulus is expected imminently, is likely to reveal areas that are anticipated as important through experience. Initial gaze was used to indicate where users expected to find target links.

Entry point: where the participant looked within 250 msec of the web page appearing on the screen. Again following Loftus and Mackworth's conclusions

about the significance of the location of initial fixation (1978), entry point was used to identify initial design elements that attracted attention.

Time to target fixation: the time it takes the participant to fixate on a specific target. Jacob and Karn (2003) suggest the metric of *time to target fixation* as a useful measure when there is a specific search task. Time to target fixation was used here to indicate visual search performance.

Gaze duration on look zones or AOIs: the amount of time the participant fixates on points within a given look zone. Fitts, Jones and Milton (1950) distinguish between gaze duration, reflecting difficulty of information extraction, and frequency of gazes, reflecting the importance of that area of display. Gaze duration on *look zones* was used to indicate difficulty of information extraction.

Transition of fixations between design elements: the number of transitions between two specified areas of interest. Jacob and Karn (2003) suggest that number of transitions between related design elements indicates the efficiency of the arrangement of elements on the user interface, with fewer transitions reflecting a more efficient arrangement, and high numbers of transitions indicating a referencing or interpretation difficulty. Transition of fixations was used to indicate efficiency of layout.

Table 4-1 presents these additional eye movement-based metrics as used in previous studies and as applied across the four research studies.

Metrics	As used in previous research	As used in this research
Scan path efficiency	Rayner and Pollastek (1994)	Study **1**
Location of first fixations: (adapted as initial gaze and entry point)	Loftus and Mackworth (1978) Byrne *et al.* (1999)	Studies **1,2,4**
Time to target fixation	Ellis *et al.* (1998)	Studies **2,3,4**
Gaze duration on look zones	Fitts *et al.* (1950), Harris and Christhilf (1980), Hendrickson (1989), Benel *et al.* (1991), Flemisch and Onken (2000)	Studies **3**
Transition of fixations between design elements	Hendrickson (1989), Kotval and Goldberg (1998)	Study **3**

Table 4-1 - Eye movement-based metrics used in other studies and in this research

4.2.4 Post-session interview

The post-session semi-structured interview used open questions about users', preferences and dislikes regarding web pages. Moreover, participants were presented with print-outs of the previously seen web pages (during the eye tracking session) in order to assist them in commenting on the overall design layout. Typical questions asked participants where they had seen specific design elements placed on other websites and where would they like these elements on an ideal site. Examples of the post-session protocols for all studies are found in Appendix A.

4.3 The stimuli

Websites are visually complex, involving many different design elements. Hence, although it might be desirable, it can be difficult to attribute performance or behaviour variations to a particular design element when evaluating usability. The use of eye tracking as an indicator of the usability of websites has received criticism (Jacob and Karn, 2003) because of the difficulty in interpreting results given the

large number of variables on a web page per study. Each study reported in the following chapters aims to focus on a limited number of specific attributes only in order to uncover usability problems and recommend further refinements.

All the stimuli were taken from websites active at the time of testing. In order to control the number of variables, it was decided to use static web pages, and certain effects were manipulated using design tools such as Photoshop, Paint Pro and Front Page. Variables such as the number of images appearing on the screen, the colour of design elements such as menu items, the position of design elements such as the 'About Us' link and the shopping cart were transformed per study design. This strategy restricted the number of independent variables and enabled the testing of each study's hypotheses.

The selected stimuli came mainly from the e-travel and e-tail domains. They were selected to ensure that:

- They were not leading websites, in order to decrease the likelihood that the participants would not have seen them before;
- They shared a similar overall design layout, so that layout was not confounded with presentation format;
- They could be modified to suit the experiment design; and
- They would not require any scrolling.

4.4 Devices

The studies conducted in this research used two video-based eye-tracking systems that project harmless infrared light onto the eyeball and track reflections off landmarks of the eye, recorded by video cameras fitted with filters. Image-processing software is used to identify and map eye position to the display. The lens, cornea, and other parts of the eye absorb a small amount of energy from the infrared light, but it is less than 1 per cent of the Maximum Permissible Exposure Level as certified by the American Standards Institute (ANSI Z 136.1-1973). This is the equivalent of light one can get on a sunny day.

There are numerous video-based systems. Study 1 used a SensoriMotoric Instruments (SMI)'s head-mounted system (HED-II) at a sampling rate of 50 Hz (SensoriMotric, 1999) that corrects for head movement and allows for general movement (Figure 4-4). Studies 2, 3 and 4 used an ASL 504 (Applied Science Laboratories) remote pan-tilt camera at a sampling rate of 50 Hz for which hardware does not come in contact with the eye or head (Figure 4.5).

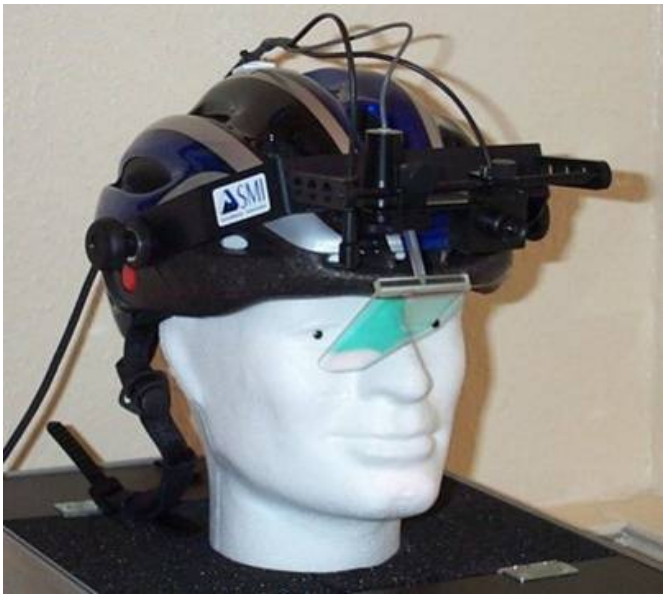


Figure 4-4 - SMI head-mounted eye tracker

Figure 4-5 - ASL 504 pan-tilt eye tracker

Systems differ in the amount of time required to set-up and adjust the hardware. In some cases additional items are used such as chairs that can be adjusted to change the height of the display relative to the participant, or **chin rests** to reduce head movements. A study by Nevalainen and Sajaniemi (2004) compared three eye tracking devices to explore their ease of use and accuracy. They also observed the

trade-offs of using each device in terms of inconvenience the participants faced. The three devices they examined were: Tobii 1750™ from Tobii Technology, ASL 504 from ASL, and the ASL 501 from ASL. They found that the ASL 501 required approximately twice as much time for the preparation as the other two devices. Tobii 1750™ seemed to be the most unobtrusive for the participants, and the ASL 504 needed to be checked manually for possible loss of data.

Both eye tracking devices used in this research were located in the Faculty of Education and Language Studies at the Open University. Study 1 used the SMI device as that was available at the time, and thereafter studies were conducted with the ASL 504 device as that device was added to the resources shortly after Study 1. The ASL 504 device was preferred, as it required less time to set-up, it was less intrusive for the participants because it did not require any physical contact with them, and the organisation of the data was eased by the use of accompanying analysis software (GazeTracker™).

4.4.1 Data collection for Study 1

The SMI's Head Mounted Eye Tracking Device (HED-II) uses two small cameras (the eye camera and the scene camera) and an infrared light mounted on a bicycle helmet, only 450g. (Figure 4-4). A piece of glass is placed on the helmet which reflects the downward pointing light into the eye. The front of the eye is illuminated with the infrared light. This produces a **bright pupil** and a **corneal reflection** (SensoriMotoric Instruments, 1999). The scene camera records what the participant is viewing on the screen. This recording and the eye movement data are transferred via video encoding hardware to the researcher's machine. This enables the creation of mpeg video files of the scene video with an over-laid moving dot

representing the participant's eye movements. The iview™ software on the researcher's computer is connected to the eye tracker to analyze the transmitted image; the Point of Regard is then computed (SensoriMotoric Instruments, 1999). The eye movement data can be exported to Excel spreadsheets for further analysis.

4.4.2 Data collection for studies 2, 3, and 4

Eye movements were recorded using an ASL 504 eye tracking remote pan-tilt camera capturing eye movement data at a sample rate of 50Hz. The camera was placed in front of the participant's screen (17 inches flat PC screen) without obscuring the participants' view. The presentation of the stimuli was controlled by means of the GazeTracker™ software and presented on the screen viewed by participants from a distance of 55 cm from the screen. The GazeTracker™ software also records the eye-movements and enables the researcher to view the data and output statistics relating to researcher-definable regions of interest on the web page. ASL Eyepos software also recorded the data and was used as a back-up. Each of the studies conducted as part of this research was technology-intensive. As the lab was shared with other members of the university this required a separate set-up each time a participant was about to come in for a session as part of the conducted studies. A detailed presentation of the set-up is presented in Figure 4-6.

4.5 Ethical considerations

It is of the utmost importance that research involving human participants is carried out to the highest possible ethical standards. The key principles taken for all four studies in this research were:

- not to harm,

- to ask for consent (an example of the consent form used for this research can be found in Appendix A),
- to allow freedom to withdraw at any time, and
- to ensure confidentiality.

These are the main principles identified by the British Psychological Society (BPS) in its ethical code of conduct (BPS, 2002).

All participants were over 18 and volunteered to take part in the studies. Information about the eye tracking device and how it operates was given during the recruitment process. Briefing of what the participant should expect and a clear account of the purposes of the research was given upon arrival to the lab. All necessary steps were taken to ensure that the participant felt comfortable and entitled to withdraw at any stage of the data collection session. The results of the studies have been accessed only by research team members as stated on the consent form. Examples of the participants' data such as quotations or scan paths have been anonymised. Upon completion of the study, the participants were asked to report any negative experience and provide feedback for further improvement of the data collection session.

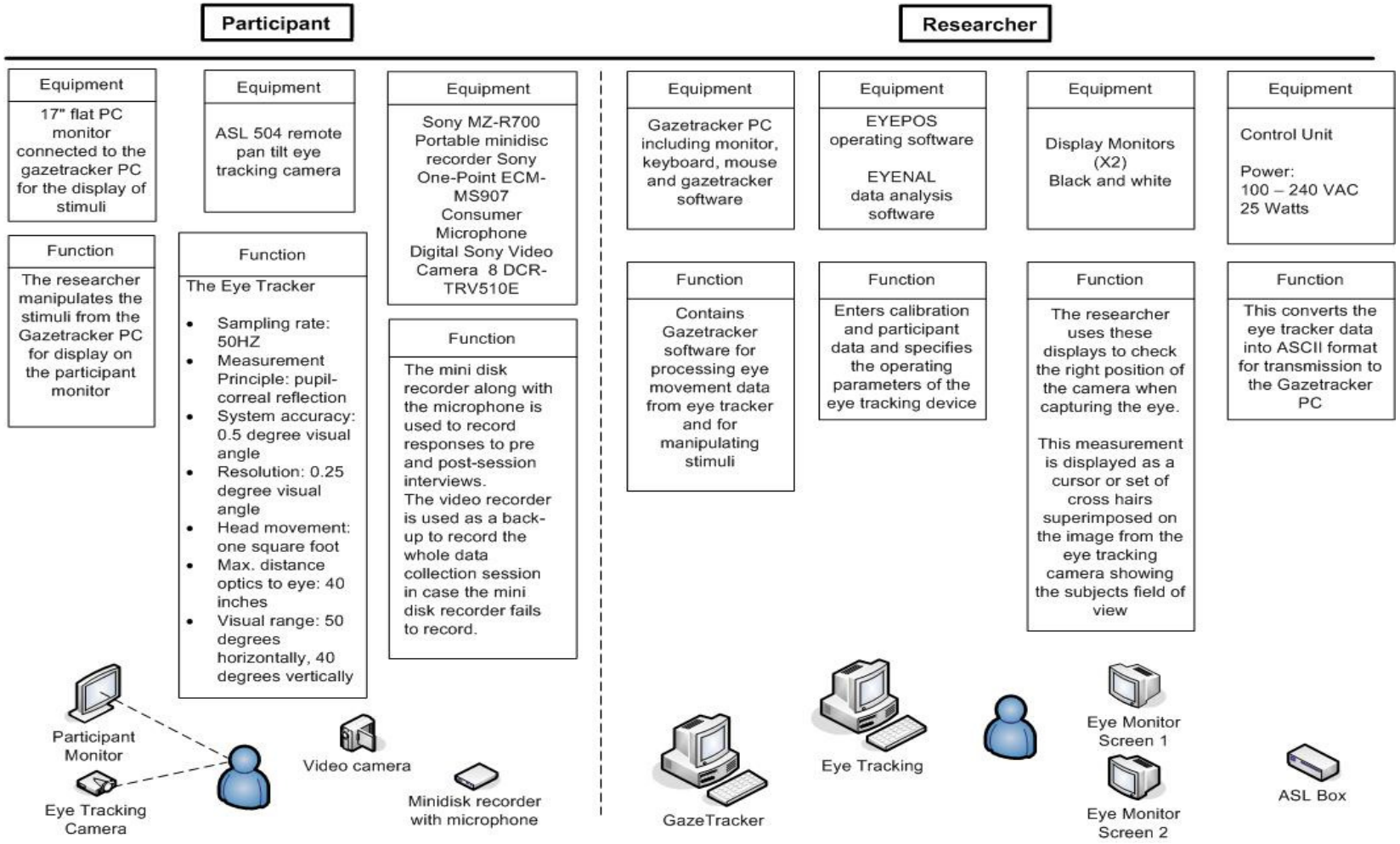


Figure 4-6 - Overview of lab set-up for data collection of studies 2, 3, 4.

4.6 Limitations

While the studies reported here were constructed carefully to overcome several limitations of earlier studies, there were still some limitations that had to be accepted. Some of these result from a limitation of resources; others are consequences of unavoidable trade-off decisions. Examples of the first type include the choice of equipment (especially the use of head-mounted equipment in study 1) and sample size (especially in study 1 and study 2) which was constrained by laboratory access and operational overheads. Examples of the latter type are discussed below.

Laboratory studies were used as they allow a high level of experimental control while keeping resource requirements relatively low. This decision carries a cost in terms of reduced ecological validity. For example, were the participants' performances during the studies similar to that in real on-line interaction? The main limitation of the reported empirical work is that none of the participants was asked to interact within a website, nor were they observed during longer tasks. They were asked only to look at single web pages, such as the homepage of the website, whereas in naturally-occurring interactions, such as using the internet in the work place, the user would have been able to access further web pages within a website. These limitations on the observed interactions might have led to limited or atypical behaviour, and this is a matter for further study. However, the experiment tasks were representative sub-tasks of authentic interaction (e.g., identifying what sort of site one has found), the sites were authentic commercial sites, and the subjects were readily able to relate the experiment tasks to their prior experience, and so there is

reason to believe that the design preserved reasonable ecological validity while allowing necessary control.

Another question is to what extent the results of the reported research can be generalised. The participants were a relatively small, volunteer sample, not a representative selection. Limitations of the eye tracking devices (which rely on reflected light) mean that spectacles-wearers were excluded from participation. As frequent or regular internet users comfortable and familiar with this technology, the participants arguably represent a significant sub-section of e-commerce customers, and cautious generalisations can be extended to that sub-section. However, they certainly do not represent the whole range of people attracted by internet shopping, and the results cannot be generalised more widely. Although the conclusions are thus limited, the studies certainly yielded well-founded preliminary insights that provide a good basis for further investigation of factors that likely influence visual search behaviour when interacting with e-commerce sites.

Chapter 5 Study 1- Homepage Study

The infinite has no beginning, but it seems to be the beginning of other things, and to surround all things and guide all…and it is the beginning which is the most important part of the work.
Epictetus (55 - 135 CE)

This chapter reports on the conduct and findings of the first in a series of eye tracking studies intended to re-explore HCI design guidelines from a finely-grained behavioural perspective, as explained in Chapters 1 and 2.

The first of these studies explored homepage guidelines. The homepage is often the initial page of a website users see. First impressions matter in terms of attracting new customers and also retaining existing ones. The homepage Study allowed the empirical investigation of existing homepage design guidelines and identified potential refinements. The first section of this chapter addresses the rationale of the study.

5.1 Rationale of Study

Existing design guidelines do not focus specifically on e-commerce websites and only one (Nielsen and Tahir, 2002) has focused on design guidelines for homepages.

The homepage is the most visited page of an e-commerce website and the first opportunity to attract potential customers. The role of the homepage of an e-commerce website is to 'anchor' the site: to identify and introduce the company and its services and hence to attract and retain customers (Nielsen and Tahir, 2002).

When arriving at an e-commerce homepage, the customer should be able to quickly recognize the purpose of the site and organisation's brand identity (Nielsen and Tahir, 2002).

The brand identity is one of the major factors that have an impact on the customers' trust in an e-commerce website (Fang and Salvendy, 2003). A good brand name gives a good first impression and evokes positive associations with the brand, such as the attributes that customers think of when they hear or see the brand name. Customers state that they trust and go to online stores with well-established brand names. Overall, if a website is considered valuable it will be remembered as one to use again or even to recommend to others (Fang and Salvendy, 2003).

Brand identity and the site's purpose should be established in the first inspection of the homepage (Van Duyne *et al*. 2003). The most recognizable feature of a brand is the name, logo, symbol or trademark that denotes a product's origin (Pavitt, 2001, p.21). Hence, the location and prominence of the logo play a role in homepage design. However, Nielsen's 'Alertbox' on the 'useit.com' site (Nielsen, 2003) suggests that one of the ten most-violated homepage design guidelines is *'emphasize what your site offers'*. Nielsen reports that only 27 per cent of homepages explicitly state what they offer users (Nielsen, 2003). The process of identifying the company and the brand takes place during the first few seconds of visiting the homepage of an e-commerce website (Van Duyne *et al.* 2003). During these first few seconds of customers glancing at the website, the homepage should communicate to customers where they are, what the company does, and the tasks the customers can perform on the website. For example, the location of the first fixation on a homepage indicates where the user expects to find information about the company. Russell (2005) found that first fixations

concentrate on distinctive design elements such as pictures and logos. If, for example, the first fixation focuses on a company logo, then brand identification is facilitated. On the other hand, if the first fixation finds a non branding-related design element, the user might be misled about the nature of the site. Hence, there is a need to examine the way customers look for specific design elements such as the logo or the company's name on homepages. By capturing the customers' eye movements when they visit a homepage for the first time, it is possible to identify which location on the homepage attracts immediate attention, and how the homepage tends to be scanned. This can be related to what information is presented in those locations (and whether immediate attention is drawn to important information), and hence it can be assessed whether the homepage conveys a clear message about the brand and site identity.

5.2 Method

Study 1 focused primarily on users' visual search such as where they fixate before a homepage appears on the screen, first fixation on the homepage indicates very early scanning behaviour while they look for the website's brand identity and main purpose. The study also considered factors that might influence the customers' expectations when searching for this information, such as previous experience visiting e-commerce homepages.

Homepage Study		
Phase 1: Background Questionnaire • User profile • Previous internet experience	**Phase 2: Eye tracking session** • Position of first fixation within given time • Reading and analysing of scan paths • Accuracy of task responses	**Phase 3: Post–session Interview** • Perceptions and expectations of information on Homepages • Interaction with print-outs of eye tracking session

Table 5-1 - Data collected during Homepage Study

The Homepage Study collected four sets of data (table 5-1):

- Background data

The background questionnaire elicited a user profile (age and gender), previous internet experience (frequency and purpose of internet use), and three examples each of familiar and frequently-used websites. The background questionnaire provided data to address how previous experience of visiting e-commerce websites influences users' expectations about where to find the name and logo of the company on homepages.

- Eye movement data

Three forms of eye movement data were recorded:

- Location of *initial gaze,* to indicate where on the homepage users *expected* to find the brand and site identity,

- *Entry point*, to indicate which design element first attracted their visual attention,

- *Scan path*, to assess how easy or difficult it was for the users to find the information,

- Task performance

The users' task was to identify the profile of the company and the activities they could do on the website. Their responses to the task questions were recorded verbatim. The content and accuracy of their responses to the task questions were used to assess their performance of the task.

- Post - session interview

The post-session interview used semi-structured questions about users' expectations, preferences, and dislikes regarding homepages. They were asked where they thought they first looked on the homepages during the eye tracking session. The aim of the interview was to elicit the users' perceptions and expectations about branding on homepages and to compare the users' perceptions of where they first looked to their actual eye movements. The interviews were recorded verbatim.

Pilot studies

An initial pilot study, with five users, was conducted using a similar structure but with only one task question: "*What is the name of the site*?" Previous research (Loftus and Mackworth, 1978) showed that 15 seconds is sufficient to capture a complete scan path for the first inspection of a display. The initial pilot study corroborated that 15 seconds was ample for a complete scan of a homepage, but it suggested some improvements to the protocol. As a result, the study design was altered to include the revised task questions, larger screen images of the

homepages, the blank screen between homepage presentations, and the post-session interview.

The background questionnaire and post-session interview questions were piloted separately with three users to establish whether the questions were comprehensible and whether they elicited relevant responses.

A second pilot study, with one user, tested the feasibility of the revised protocol and the effectiveness of the background questionnaire and post-session interview questions. The pilot study again corroborated that 15 seconds was sufficient for a complete scan of a homepage and that the questions were effective.

5.2.1 Study design

Study 1 focused on customers' visual micro strategies – such as where they fixate before a homepage appears on the screen, first fixation on the home page, and very early scanning behaviour – while they look for the website's brand identity and main purpose. It also considered factors that might influence the customers' expectations when searching for this information, such as previous experience visiting e-commerce homepages.

The study questions were:

- Does the overall design layout of homepages influence initial inspection?
- Do previous experience and preferences influence expectations of initial inspection of homepages?

The users' task was to identify what the company did and what they could do on the website. After a briefing about the study, in particular the use of the eye tracker, each participant completed a consent form and a written background questionnaire. The participant was asked to read a pair of task questions before starting the eye tracking session:

- "*What does this company do*?"
- "*What can you do in the site*?"

The eye tracking equipment was then calibrated for the individual. The five selected homepages appeared in sequence as a Power Point™ slide show. First, a blank screen was presented until the user indicated readiness, then one homepage was presented for 15 seconds, followed by a blank screen while the participant answered the task questions for that homepage, followed by the next homepage, and so on. The order of presentation of the homepages was varied for each participant, in order to reduce possible order effects. The task questions were the same for each homepage. The participant's eye movements (fixations and saccades) were recorded before and during inspection of each homepage. The participant's spoken responses were audio recorded.

A post-session, semi-structured interview followed. During the interview, participants were shown print-outs of the homepages, so that they could indicate on each point where they thought they had looked first. They were also asked to comment on the design layouts.

5.2.2 Participants

Ten participants (five male and five female) were recruited within the postgraduate students and staff of the Open University. Their ages ranged from 23-51. All participants were regular internet users (using the internet 2-3 times per day) and two classified themselves as frequent internet users (using the internet throughout the day as part of their job). None of the participants had viewed the homepages used in the study prior to their participation.

5.2.3 Stimuli and equipment

Five e-commerce homepages were selected as stimuli for this study:

Disney www.disney.com (entertainment),

Petsmart www.petsmart.com (pet supplies)

General Motors www.gm.com (automotive)

Federal Highway Administration www.fhwa.dot.gov (transportation), and

Global Sources www.globalsources.com (office supplies).

See Figures 5-1 to 5-5.

Figure 5-1 - *Disney* homepage

Figure 5-2 - *Federal Highway* homepage

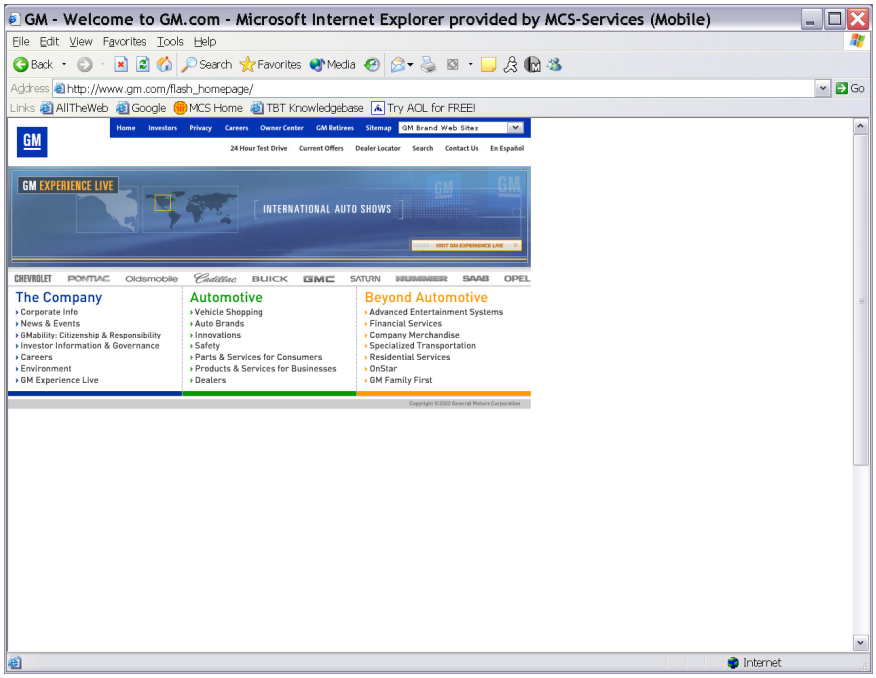


Figure 5-3 - *General Motors* (GM) homepage

Figure 5-4 - *Global Sources* homepage

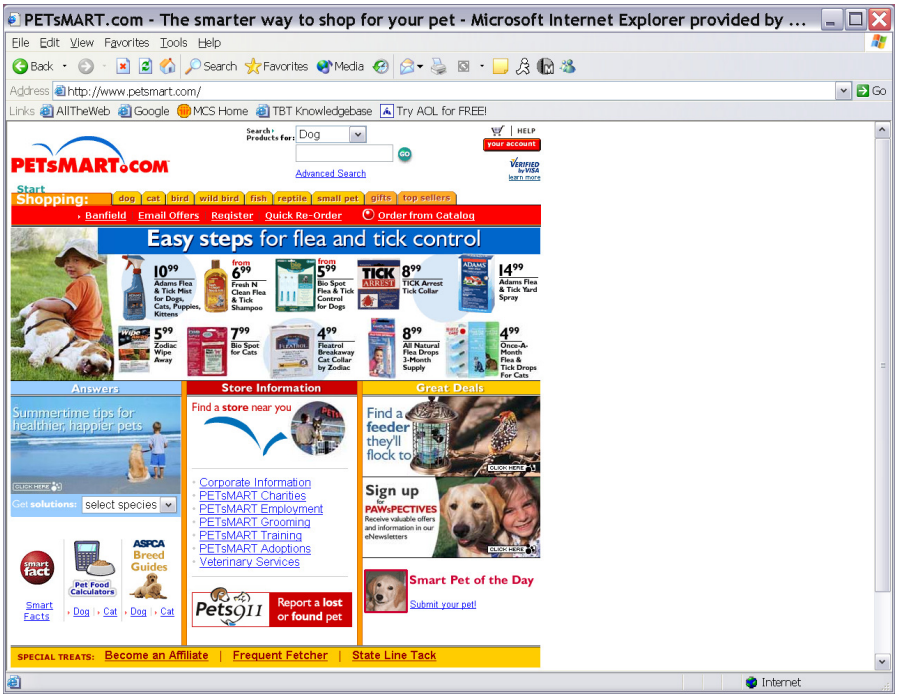


Figure 5-5 - *Petsmart* homepage

The five were chosen from the 50 homepages evaluated as part of a study by Nielsen and Tahir, (2002). The aim of the study reported here was to compare the design guidelines derived by (Nielsen and Tahir, 2002) to the outcomes of our eye tracking study. Therefore, the study used homepages cited by Nielsen and Tahir and which retained the same design layout since the heuristic evaluation by Nielsen and Tahir (2002).

Eye movements were recorded using a SensoriMotoric Instruments (SMI)'s Head-Mounted Eye Tracking Device II (HED-II) at a sample rate of 50 Hz (SensoriMotoric, 1999), as introduced in Chapter 4. The head-mounted camera captures the images of the participant's eye and scene of view. It tracks the eye by the detection of two reflections of a low intensity beam of **infra red light** from the eye tracker: one from the participant's retina and the other one from the cornea. The computed gaze position is overlaid on the scene image and visualized in real time.

An integrated MPEG video file was produced representing the participant's eye movements.

5.2.4 Procedure

After a brief introduction about the study, each participant completed a consent form and a background questionnaire. The eye tracking equipment was then calibrated for the participant. The five selected homepages appeared as a slide show. The participants were asked to answer two task questions.

The participant's eye movements were recorded, and the participant's verbal responses and interactions with the web pages were audio- and video-recorded. A post-session, semi-structured interview followed which, was also audio-recorded.

5.3 Data analysis

The analysis of the data for Study 1 focused on initial impressions of homepages. The questions in this study looked at the initial scanning of homepages. The eye movement-based metrics used to answer the study questions were: *initial gaze*, *entry point* and *scan path efficiency*.

The data from the other techniques used enabled the investigation of possible relationships between the users' visual behaviour and task performance. When interpreting previous experience, the examples of previously used e-commerce websites the users' had given in the background questionnaires were taken as representative examples of websites with which the users were familiar. The design layout was examined and then compared with the stimuli homepages. Since the

main focus of Study 1 was first impressions of homepages when looking for the brand identity, specific features were compared such as the location of the logo and the company's name on the stimuli and previously used homepages. For example, one participant gave examples of websites he uses every day where the logo was placed on the top left corner. When examining initial gaze as an indication of expectations, it was possible to see that the first area he looked at on the screen was the top left corner which hinted that he expected to find the logo in this position.

Moreover, the eye movement data were supported with data from the questionnaire and interview to explore how previous experience and preferences might effect initial scanning of homepages. When analysing task performance in terms of accuracy of responses to the task questions, four categories of responses were used:

- *Correct* when the participant was able to identify what the company does and what they can do on the site
- *Incorrect* when the participant was not able to convey the company's and site's purpose,
- *Uncertain* when the participant was not sure about what the company does and what they could do in the site, and
- *Selective* when the participant gave partially correct answers regarding the brand and site identity.

5.4 Results

The four types of data were analysed and compared. The qualitative data added value to the performance and eye movement data by revealing something of the users' expectations and perceptions. The overall outcomes of the eye tracking study were compared to Nielsen and Tahir's (2002) homepage design guidelines.

Eye Tracking Session

During the eye tracking session each participant's scan path was recorded (Figure 5-6). Two measures were used:

- 'initial gaze': where the participant looked 100 msec prior to the homepage appearing on the screen, and
- 'entry point': the first fixation within 300 msec of the display of the homepage. 300 msec is the duration of a typical fixation and the typical time in which information extraction occurs (Cowen *et al*. 2002).

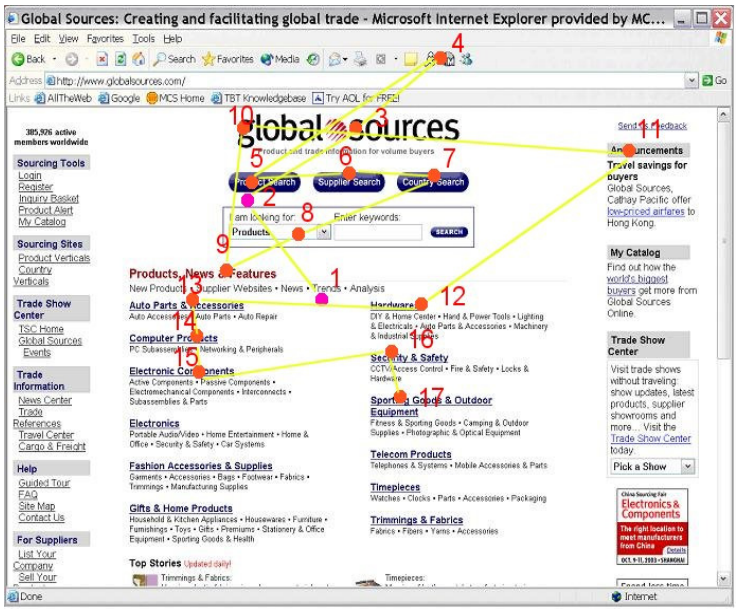


Figure 5-6 - Sample scan path on homepage

A common trend was found for the eight users who classified themselves as 'regular internet users'. Their 'initial gaze' was in the middle of the screen across all homepages (on or near the boundary between areas B and E in Figure 5-7). The two participants who classified themselves as 'frequent internet users' looked at the top left corner (area A) or top middle (area B) of the page. Although the small sample means that no firm conclusion can be drawn, this observation warrants further investigation.

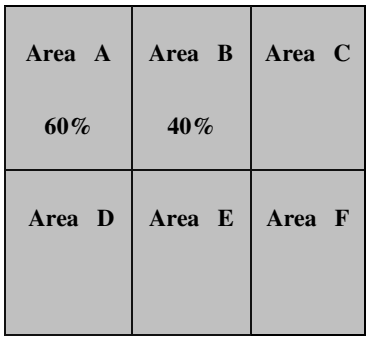


Figure 5-7 - 'Entry point' distribution across all homepages

Figure 5-7 shows the 'entry points', or first fixations when the homepage appeared on the screen. No distinction was evident between the 'regular internet users' and 'frequent internet users'. Figure 5-8 shows the frequency of response types across all homepages for Task 1 and Figure 5-9 for Task 2.

The *GM* homepage gathered the highest proportion of correct responses in both task questions, that is, the site best enabled users to identify what the company does and what can be done within the site. The *Disney* and *Petsmart* homepages were the only homepages to evoke incorrect answers.

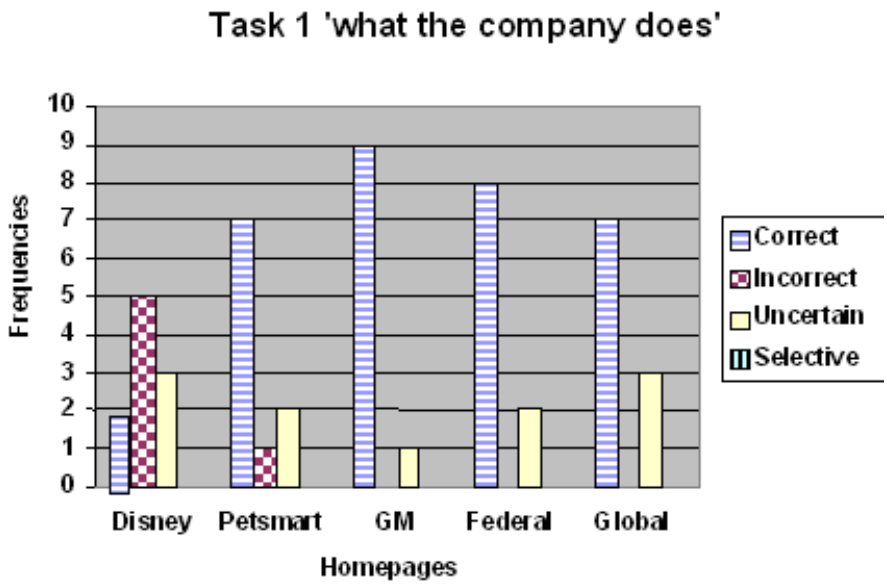


Figure 5-8 - Frequency of responses for Task 1 '*What does the company do'* across all homepages.

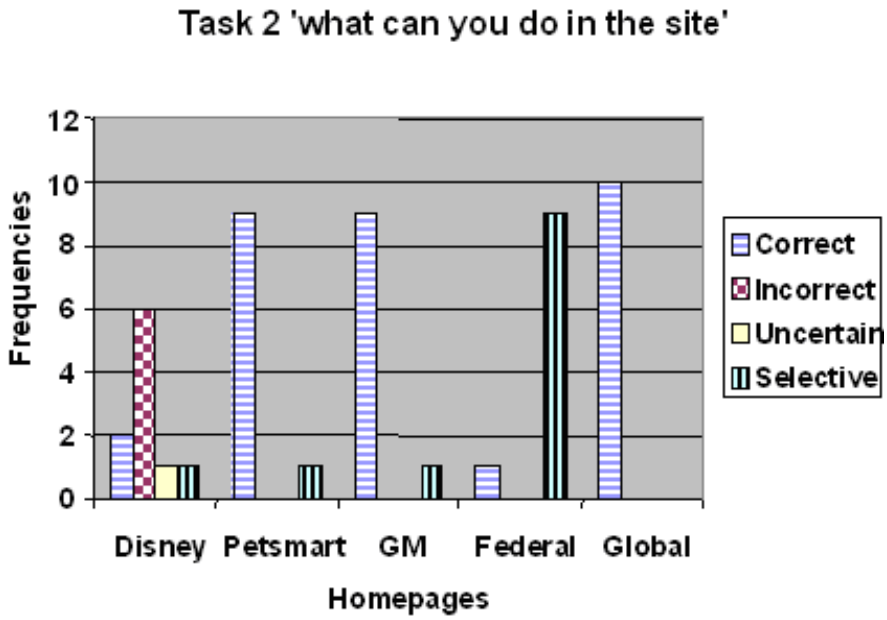


Figure 5-9 - Frequency of responses for Task 2 *'What can you do in the site'* across all homepages.

Background Questionnaire and Post-session Interview

Based on the completed background questionnaires, the websites the participants visited regularly were investigated. All placed the logo and name of the company on the top of the homepage, either the top left corner or top middle. This placement corresponds to the 'entry point' fixations recorded for the participants.

The participants' self-reports on where they first looked on each homepage were compared to the 'entry point' data. Users were uniformly accurate in their reports. They knew where they looked first. The participants' comments about design layout were subjected to inductive analysis. Emergent topics, and the design characteristics users associated with them, are summarized in Table 5-2.

Emergent topics	Associated design characteristics	
Annoying/frustrating design elements	• Too much text • Product Images • Adverts	• Introduction page • Pop up windows • Small fonts
General preferred design elements	• Simple/clear links • Images • Sub links • Less text • Company's contact details	• Short description under bold titles • 'About the company' link • Clear structure • Easy access to products
Expectations of homepages	• The name of the company/logo • Links of the rest of the site • The company info • Large fonts of company name	• Title of page • Generic information • Navigation tool • Keywords/phrases
Perceived position of 'entry point'	• Top of the page top middle or top left corner	• Left side of the page
Perceived first design element to look for	• A peripheral look to confirm it is the site aimed for	• The name of the company • Links to the rest of the sites
Perceived factors that influence first fixation	• Information presented on paper documents • Left to right reading	• Combination of previous experience and cultural aspects • Visiting other websites

Table 5-2 - Post-session responses

5.5 Discussion

Disney Homepage

The *Disney* homepage consists of a centred design layout with the core of the homepage presented in the centre of the page as defined by Bernard and Larsen (2001). On the top of the page there is a large advertisement banner about the *Disney* Visa Card. Figure 5-10 demonstrates a representative scan path of a user looking at the *Disney* homepage. The dots represent fixations. The *Disney* homepage caused confusion among the participants regarding the brand identity. During the eye tracking session all participants' entry points appeared on the top middle of the homepage were the advertisement banner was positioned. But two participants' initial gazes before the homepage was presented appeared on the top middle of the page as well, whereas the other eight participants' initial gazes appeared in the middle of the page. The two participants' with top middle page initial gazes had classified themselves as 'frequent internet' users. This suggests a difference of initial gazes between the 'regular and frequent internet' users; however this can only be a speculation without further investigation. The fact that the entry points of all participants were on the top middle of the page where the advertisement banner was suggests that the user confused the banner with the brand as the users looked where they expected to find information about the company. The task responses reflect this confusion as well. Only two participants were able to identify the brand of the company whereas five of the participants mistook the advertisement banner on the top of the page for the main purpose of the site. All the

participants' scan paths reached the bottom of the page suggesting complete scan paths of the context of the page. One of the participants typically said during the task completion: *'I'm not sure if the company is Disney Visa Card or Disney. This is strange and confusing, what is the site about?'* The site's identity was reported as even more confusing as six participants incorrectly responded that they could buy products online or book events for the theme park with a *Disney* Visa Card instead of responding that they could apply for a Visa *Disney* Card online.

Figure 5-10 - Example scan path of user viewing *Disney* Homepage

Petsmart Homepage

The *Petsmart* homepage consists of a left justified layout as defined by Bernard and Larsen (2001) with the core of the homepage appearing on the left of the page remaining blank (white space). The homepage's characteristic feature is that it includes many images of *Petsmart*'s products (Figure 5-11). During the eye tracking session all participants' 'initial gazes' fixated on the middle of the page before the homepage was presented on the screen and their entry points were on the top left corner of the homepage where the name and logo of the homepage was positioned.

The users' expectations of finding the brand of the company on the top left corner of the page was met since the users first looked at that position. In addition to the match of expectations, the scan path length of all participants suggest effective and complete scanning, because it reached the bottom of the page. In addition to the eye movement data, the task responses suggest a clear understanding of what the company does and what can be done in the site. Seven participants correctly reported that the company sells products for pets; only one participant was not able to report the site's brand identity. In addition, nine participants correctly identified the site's main purpose.

During the post session interview when the participants were asked to express any preference or dislikes about the selected homepages, many users expressed their dislike about the images of the products being displayed on the homepage. For example, '*In some cases I don't like images, here the Petsmart site does a familiarity thing, something like a supermarket, but I don't like it, I think it is*

aesthetically horrible'. Nevertheless from the eye tracking session it was revealed that despite participants not liking the images their attention was caught by them.

Figure 5-11 -Example scan path of user viewing *Petsmart* Homepage

General Motor Homepage

The *General Motor* homepage consists of a design layout that appears on the top left corner of the page. The resolution of the screen the homepage was presented on during the eye tracking study was 1280 by 1024 **pixels**. Figure 5-12 represents an example scan path of a participant viewing the General Homepage. Eight of the participants' 'initial gaze' was on the middle of the screen prior to the homepage being presented on the screen, whereas two of the participants' 'initial gaze' was on the top left of the screen. All ten participants' 'entry points' were on the top left corner of the screen where the logo of the company is positioned. The participants' scan path length suggests a complete scanning of the homepage. This is also

reflected by the high frequency of scanning of correct task responses. Nine participants were able to identify both brand and site identity. Only one participant was uncertain and selective in his/her responses.

During the post-session interview some of the participants expressed negative comments about the presentation of the design layout on the screen. One typical quote was: '*This layout is confusing, it's small just in one corner of the page, and I'd prefer it bigger!'* Such comments are contradicted by the eye movement data which suggest that it is easier to scan design layouts that appear on a small proportion of the screen. In addition to the eye movement data, the task responses showed that the users were able to answer the task questions correctly.

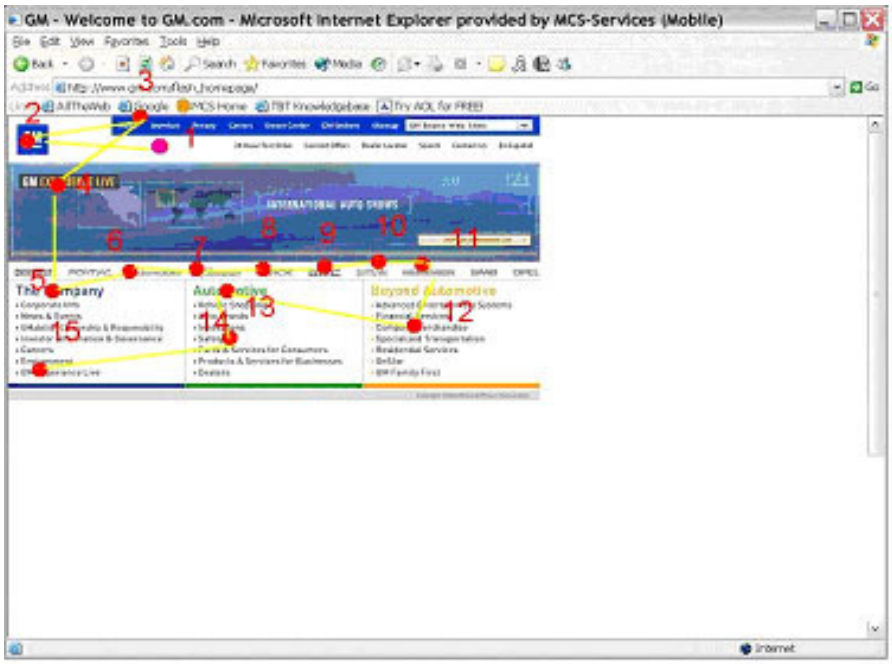


Figure 5-12 - Example scan path of user viewing *General Motor* Homepage

Federal Administration homepage

The *Federal Administration* homepage has a fluid layout of which the margins are not fixed at any particular width as defined by (Bernard and Larsen (2001). Figure 5-13 shows a typical scan path of a user that has not scanned to the bottom of the homepage. During the eye tracking session all ten participants' entry points appeared on the top left of the homepage where the logo and name of the company is positioned. The 'initial gazes' of all participants were on the middle area of the screen.

Figure 5-13 - Example scan path of user viewing *Federal* Homepage

Eight participants were able to report the brand identity of the site but two were uncertain about it. Only one participant was able to identify what he could do within the site, the other nine participants were selective in their responses without reporting the main elements of the site. Based on the participants' sequence of eye movements it was possible to see that although first fixations appeared on the top left corner of the page where the logo of the company is positioned, fixations were

drawn to design elements that stood out directly. For example, the link 'job opportunities' appears in red colour on the top right of the page. The participants' whose task-answers were selective were drawn to this link and consequently suggested it as the site's elements. In addition to this confusion all the participants' scan paths were found to be short, without reaching the bottom of the page indicating low efficiency in scanning. The participants' comments also confirmed this as they were characterized with confusion with the overload of text and the spread out information of the homepage. For example: '*I find sites like this annoying, there is too much information, and it looks cluttered'*.

Global Homepage

The *Global* homepage has a fluid layout as defined by Bernard and Larsen (2001): the design layout of the page fits within the screen (Figure 5-14).

Figure 5-14 - Example scan path of user viewing *Global* Homepage

'Initial gazes' of all ten participants appeared on the top middle of the page where the name and logo of the company appears. Scan path length was short for all participants similarly as for the *Federal* homepage. The participants did not scan down to the bottom of the page which demonstrates the difficulty of information extraction. This was confirmed by the uncertain task responses of the three participants who were not certain about the purpose of the company. But ten participants correctly answered what they can do in the site. This might be the result of the organization of the context of the design layout which includes white space between columns of text.

In addition to the high frequency of correct task responses during the post session interview the *Global* homepage received positive comments such as: '*I like the organization of this page and the 'look and feel' of the Google homepage with the search box and the logo on the top middle of the page'*.

Complementing existing homepage design guidelines

The main aim of Study 1 was to revisit existing homepage design guidelines in light of eye tracking data from users doing realistic tasks. It was anticipated that previous experience and expectations influence the visual search behaviour and performance of the users when visiting homepages for the first time.

When comparing (Nielsen and Tahir, 2002) design guidelines to the outcomes of Study 1 there is agreement on general design guidelines such as '*Place important information on top of the page*'. This was confirmed by the eye tracking study with many fixations on the top of the page. There are mismatches however for more specific guidelines. It is not to claim that the existing guidelines (Nielsen and Tahir,

2002) are untrue but to suggest that Study 1 adds the necessary clarity that would complete the picture and bring an insight for designers. Below the main findings have been classified in sections by homepage elements.

Banner Advertisements

In the homepage study conducted by (Nielsen and Tahir, 2002) it is suggested that "*users tend to ignore anything that looks like a banner ad so it is a poor way of promoting site elements"* (p. 125). Study 1 suggests that whether or not a banner ad will be ignored depends on its position and presentation format. For example, the banner on top of the *Disney* page was mistaken for the name of the company, due to its position and size. The 'entry points', task responses, and self-reports all confirm this confusion.

Product Images

It is suggested in Nielsen and Tahir (2002) that '*Petsmart's biggest strength is that it shows examples of the products and content offered on the site'* (p. 241). This study suggests that individual dislike for images on homepages may conflict with the images' advantages in 'drawing the eye'. All participants fixed on product images, even those who disliked and claimed to ignore them. However, those who disliked the use of product images gave less accurate responses about brand and services. Although this requires further investigation, the use of product images on the homepage could be an obstacle in search efficiency for those who dislike them.

Design Layout

Nielsen and Tahir (2002) argue that designers should use 'liquid layout' that allows users to adjust the homepage size. For example, the *Federal* homepage,

using a 'liquid layout' that filled the screen, was described as '*Well defined and easily recognisable*' (p. 161), whereas the General homepage, which had a fixed size and did not fill the (1280 by 1024 **pixels**) display, was described '*as one that doesn't make clear whom it is trying to serve or what users can do*' (p. 185). However, completeness of inspection and accuracy of brand and service identification in Study 1 ran contrary to the guideline. Users inspected the fixed-size GM homepage fully, whereas their inspection of the full-screen *Federal* homepage was incomplete, reaching only half way through the homepage. Further, there were more correct task responses for the GM homepage than for the *Federal* homepage.

Overall, Study 1 focused on initial impressions of e-commerce websites. Taking under consideration the importance of the homepage when deciding to 'stay' within a site it was decided that Study 1 would explore the factors that influence initial fixations in relation to expectations and preferences. Chapter 6 below discusses Study 2, the second study conducted as part of this research.

Chapter 6 Study 2 - User adaptability Study

Be not swept off your feet by the vividness of the impression, but say 'Impression, wait for me a little. Let me see what you are and what you represent. Let me try you'.
Epictetus (55 - 135 CE)

Another important usability principle is to conform to conventions within a website (Nielsen *et al.* 2001). When looking at e-commerce sites, meeting the users' expectations is crucial in order to increase the chances of attracting and retaining a customer. While Study 1 explored initial impressions of homepages Study 2 examines how quickly users adapt to an unfamiliar design layout and, in particular, how quickly they adjust their expectations of where to look for a given target link during repeated exposures to a new layout. Eye movement-based metrics such as time to target fixation, location of first fixation and scan path (sequence of fixations) were applied to capture users' eye movements. These metrics were then applied to analyse the effects of repeated exposures and of design layouts of websites.

6.1 Rationale of Study

As Jones and Dumais (1986) stated it "*it is not enough to know what we are looking for, we must also know where to look for it*" (pg.43). This is evident when users encounter web pages where they are presented with an overwhelming amount of information, with a mix of visual and textual design elements clamouring for attention. Thus, understanding the factors that influence visual search behaviour on the user interface is important so as to design usable user interfaces.

Existing design guidelines embody assumptions that visual search behaviour is shaped by expectations, hence they suggest designing user interfaces that conform to conventions. Nielsen *et al*. (2001) underline the importance of maintaining consistency with other websites and pages. "*All web pages are much the same from the user's perspective, they share interaction techniques, they are downloaded (slowly) from the internet, and they have relatively similar layouts. Those similarities are in fact good because they allow users a measure of transfer of skill from one site to the next. Users complain bitterly when a site doesn't try to use navigation from the majority of other sites*" (pg. 189). But how do users learn conventions and develop expectations? Ehret (2002) suggests that when locations of design elements remain constant, performance improves over exposures as users learn placements of design elements and focus/limit the scope of their visual search behaviour. The consistent placement of design elements influences visual search.

But a recent eye tracking study by McCarthy *et al*. (2003) investigated the impact of changing the placement of design elements and how users performed when viewing the element in unexpected placements. The authors found that following conventions with other websites did not matter, as users quickly adapted to design layouts with unexpected placement of design elements.

So is it important to be consistent with other websites or is it acceptable to place design elements in non-consistent locations on the user interface? The study reported in this chapter used an eye tracking device to capture users' eye movements and investigated how quickly users adapt to different placements of a design element over repeated exposures. Specifically, the study focused on one

design element, the ‘About Us’ link on a website, and addressed the following questions:

- Q1: Do users adapt their expectations of where to look when presented with repeated exposures of the ‘About Us’ link in a consistent, but unconventional, position?

- Q2: Do users adapt to alternative design layouts such as finding the ‘About Us’ link in an unconventional position?

- Q3: Where do users look first, after the repeated exposures, on pages that do not contain any ‘About Us’ link?

Previous research as conducted by McCarthy *et al*. (2003) had investigated whether placing design elements, such as menu items, in unexpected positions has an impact on visual search behaviour in terms of search performance. Variations of simple and complex web pages presented menu items in three different locations (top, left, right) for three different tasks. Each type of web page received three page visits. Study 2 of this research has applied a different experiment design: consistent placement of a design element over repeated exposures, but placed the design element at an unconventional position (that is, not as per the guidelines) to explore user adaptation. Moreover users were also presented with web pages that did not contain the target design element (‘About Us’ link) in order to explore users’ expectations of its location. Complementing eye movement data and **conventional techniques**, such as self-reports of expectations and preferences as elicited from pre- and post-session questions, has enabled a better understanding and interpretation of the eye movement data.

6.2 Method

Study 2 collected four types of data (Table 6-1):

- *A background questionnaire* to elicit the participants' internet experience and typical frequency of usage;
- *Pre-session questions* to elicit expectations about the target link ('About Us' link);
- *Eye tracking data* to measure visual search behaviour and performance; and
- *Post-session questions* to address perceptions and preferences about the 'About Us' link.

User-adaptability Study			
Phase 1: Background Questionnaire	**Phase 2: Pre-session Interview**	**Phase 3: Eye tracking session**	**Phase 4: Post–session Interview**
• User profile • Frequency and purpose of internet use	• Perceptions and expectations of placement and format of the 'About Us' link	• Time to target fixation • Location of first fixations • Scan path reading (sequence of fixations)	• Perceptions and preferences of the 'About Us' link • Interaction with print-outs of eye tracking session

Table 6-1 - Data collected during User-adaptability Study

6.2.1 Study design

A within-participants experimental design was used. The order of the ten exposures of web pages was counterbalanced (Campbell and Stanley, 1963). Each participant saw a total of ten different web pages. Counterbalancing of the presentation order of the stimuli was used to control against possible effects being attributed to specific web pages appearing in the same order across all participants.

Ten web pages of e-commerce sites were selected and amended so that they would appear in each of the three different exposure styles. So, for example, each web page was amended in order to have:

- The 'About Us' link at the bottom of the page,
- The 'About Us' link at the top of the page, and
- No 'About Us' link. The description of each exposure and purpose is presented in Table 6-2.

The study had two sets of hypotheses: The *first set* addressed the effect of the consistent placement of the target link ('About Us' link) for Exposures 2 to 7 and predicted that the placement of the target link at the bottom of the page over six repeated exposures would result in a decrease in the participants' time to target fixation and also cause changes in the participants' expectations of where to find the target link. The *second set* examined the effect of the alternative placement of the target link ('About Us' link) and predicted that the placement of the target link at the top of the page in Exposure 8 would result in quick adaptation to an unexpected but conventional design layout. Specifically it was anticipated that the first set of hypotheses will be evidenced by shorter times to target fixation, modifications of scan patterns, change of location of first fixations and self-report of preferences and expectations.

Exposures	Description	Purpose
Exposure 1	Did not include the 'About Us' link	To explore users' expectations of where to find the 'About Us' link before the repeated exposure session
Exposure 2-7	Six repeated exposures where the 'About Us' link appeared at the bottom of the page	To examine the effect of consistent design element placement on visual search behaviour
Exposure 8	The 'About Us' link appeared at the top of the page	To capture the users' visual reactions when introduced to an alternative design layout after being presented with the repeated exposures in which the 'About Us' link appeared at the bottom of the page
Exposure 9	The 'About Us' link appeared at the bottom of the page again	To assess persistence of any affect of repeated exposures on visual search behaviour
Exposure 10	Did not include the 'About Us' link	To explore users' expectations where to find the 'About Us' link after the repeated exposure session

Table 6-2 - Description and purpose of exposures

6.2.2 Participants

Ten volunteers (five male and five female) with age range of 22-56 from the staff and postgraduate student population of the Open University participated in Study 2. Five participants were 'regular internet' users (using the internet two to three times per day) and five participants were 'frequent internet' users (using the internet throughout the day as part of their job). Similarly as in Study 1, none of the participants had viewed the web pages used in Study 2 prior to their participation.

6.2.3 Stimuli and equipment

Prior to the selection of the stimuli, a survey was conducted to identify the position of the 'About Us' link on homepages. Fifty European and fifty US e-commerce sites were chosen on the basis of their sales. The percentage of homepages that placed the 'About Us' link on the top of the page as a global

navigation was 80 per cent; its position in the navigation bar varied on different websites. On the basis of this survey, it was concluded that the convention is to place the 'About Us' link at the top of the page. This led to the assumption that Web users will be used to finding the 'About Us' link on the top of the page or at least expect to find it around that position based on their previous experiences. Therefore, the position of the target link in the repeated exposures session (Exposures 2-7) was on the bottom of the page.

Ten UK e-commerce homepages were selected as stimuli for this study:

Cover4students www.cover4students.com (campus insurance),

Hatton Garden Online www.hattongardenonline.com (jewellery),

Diamond Daisy www.diamonddaisy.com (jewellery),

Travelodge www.travelodge.co.uk (accommodation),

Travel Bag www.travelbag.co.uk (travel),

Train Line www.thetrainline.com (travel),

Saga Holidays www.sagaholidays.com (holidays),

Hotel net www.hotelnet.co.uk (accommodation),

Past Times www.past-times.co.uk (gifts),

To Book www.tobook.com (accommodation).

Images of the stimuli web pages are shown in Figures 6-1to 6-10.

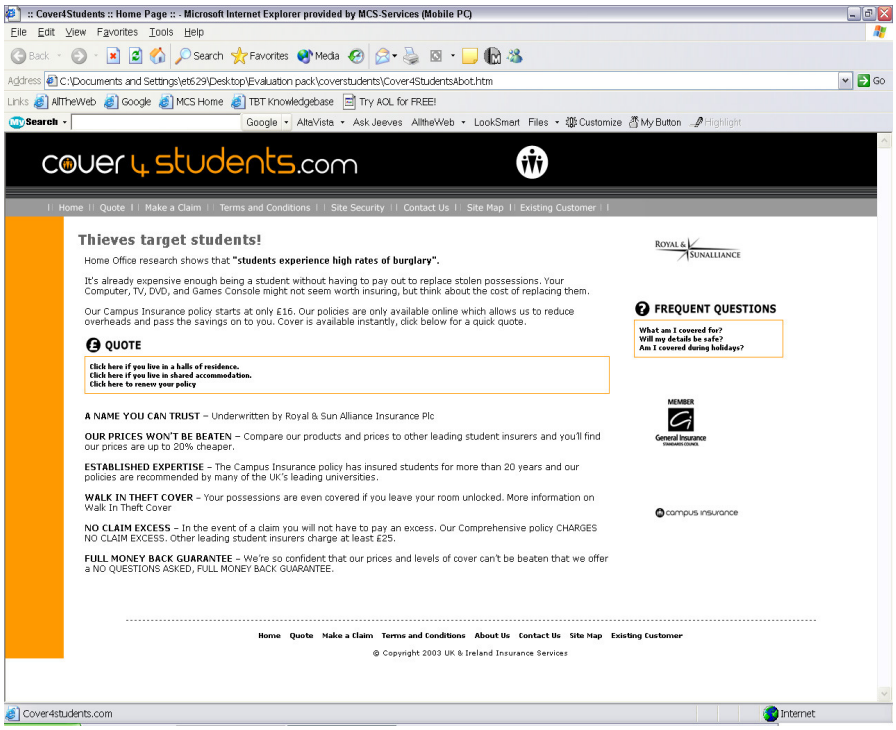


Figure 6-1 - Image of *Cover4students* web page

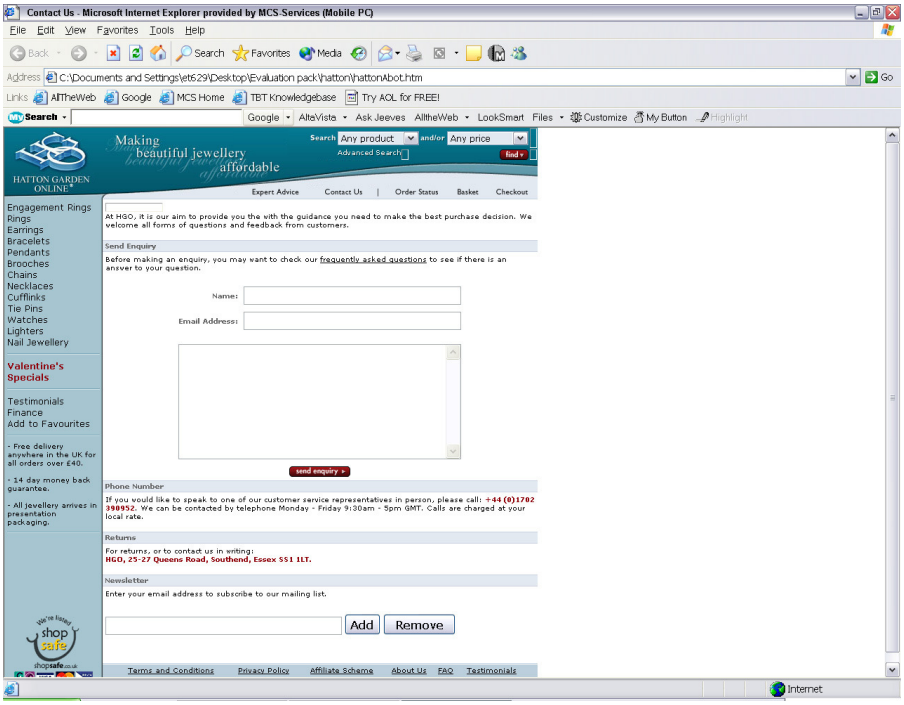


Figure 6-2 - Image of *Hatton* web page

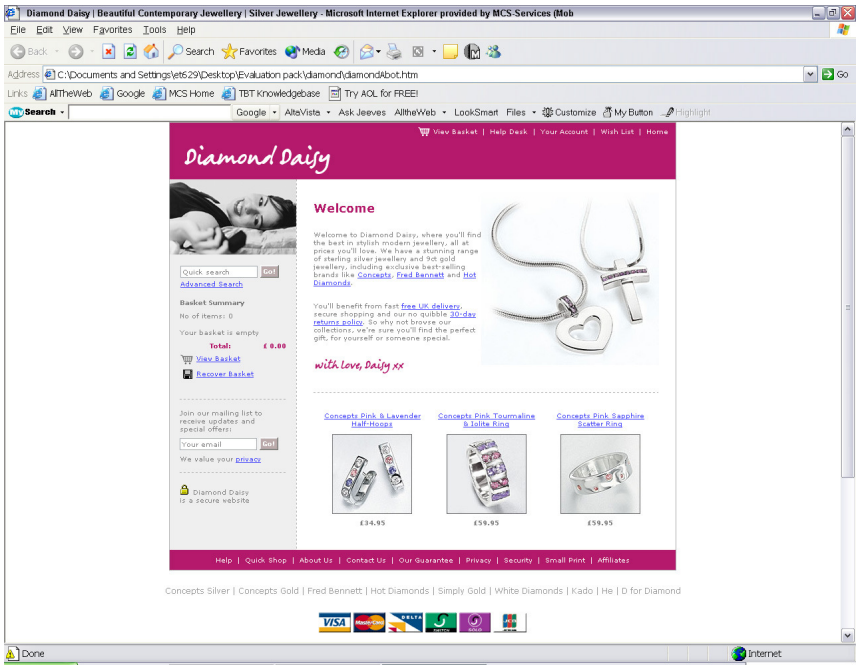


Figure 6-3 - Image of *Diamond* web page

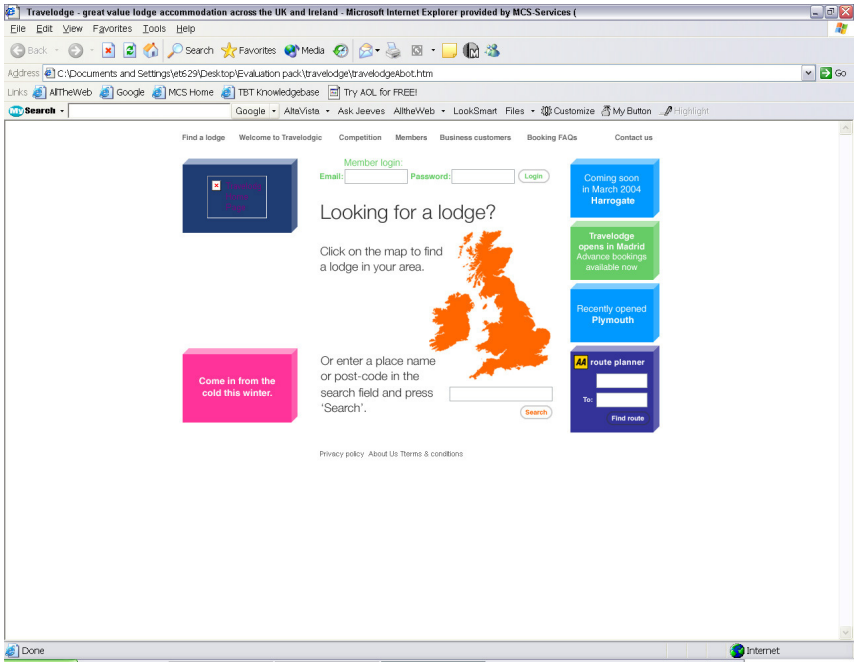


Figure 6-4 - Image of *Travelodge* web page

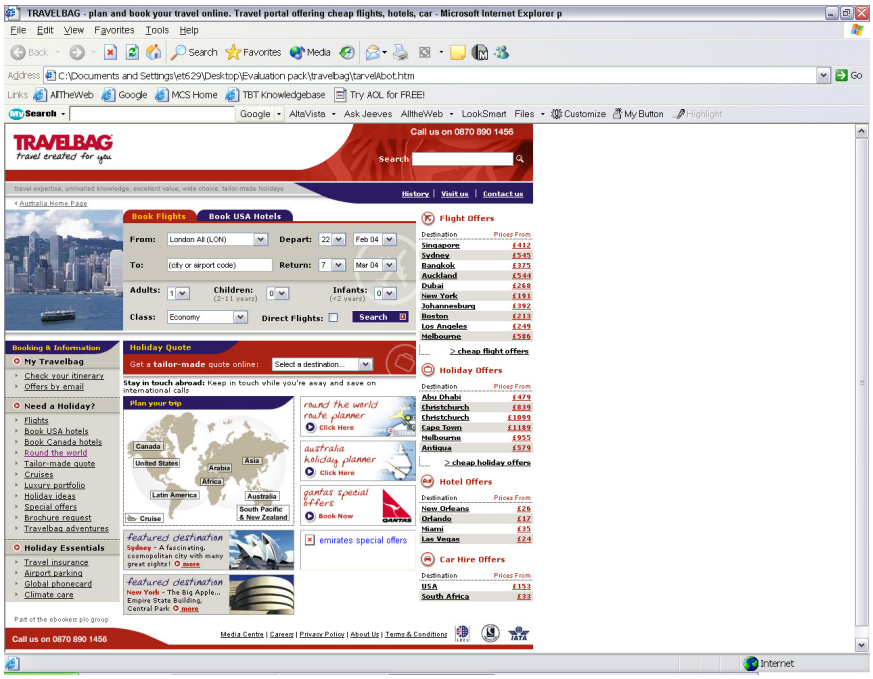


Figure 6-5 - Image of *Travelbag* web page

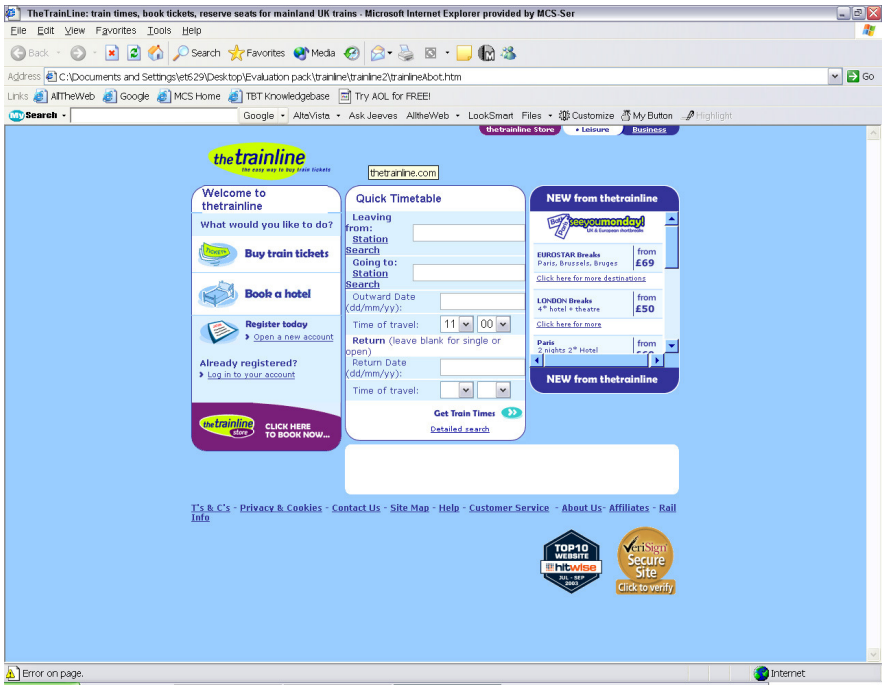


Figure 6-6 - Image of *Trainline* web page

Figure 6-7 - Image of *Saga* web page

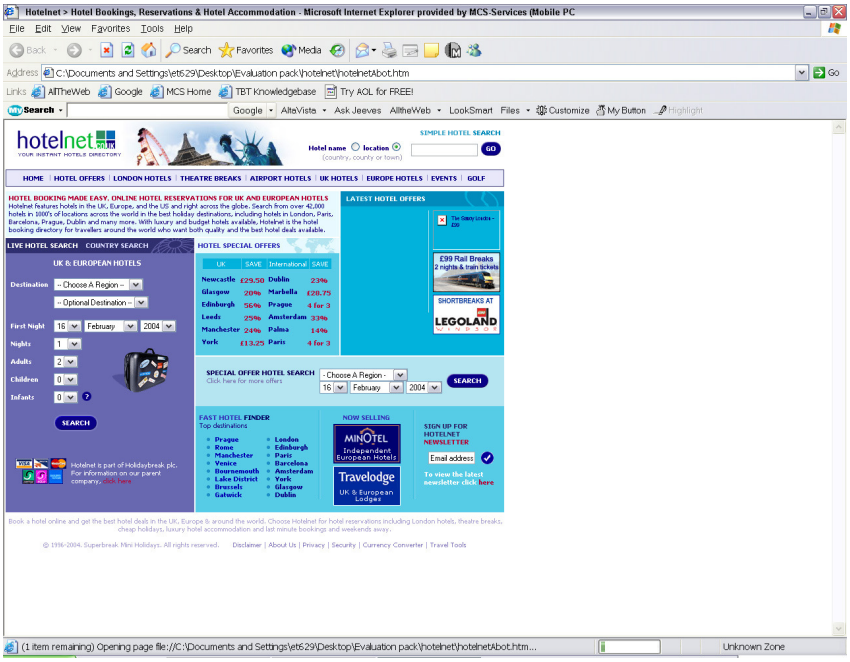


Figure 6-8 - Image of *Hotelnet* web page

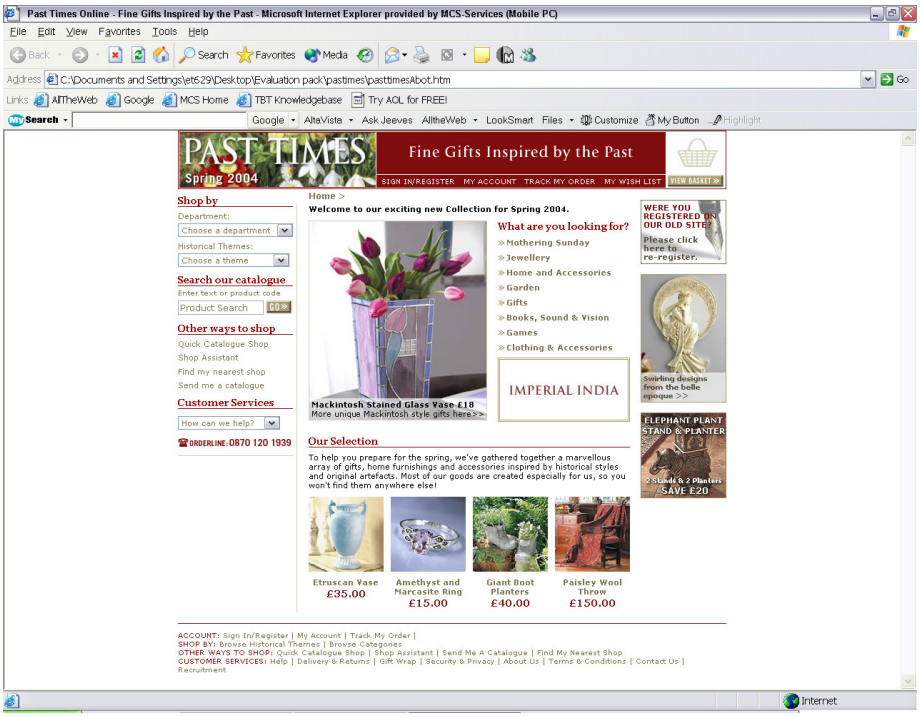


Figure 6-9 - Image of *Pasttimes* web page

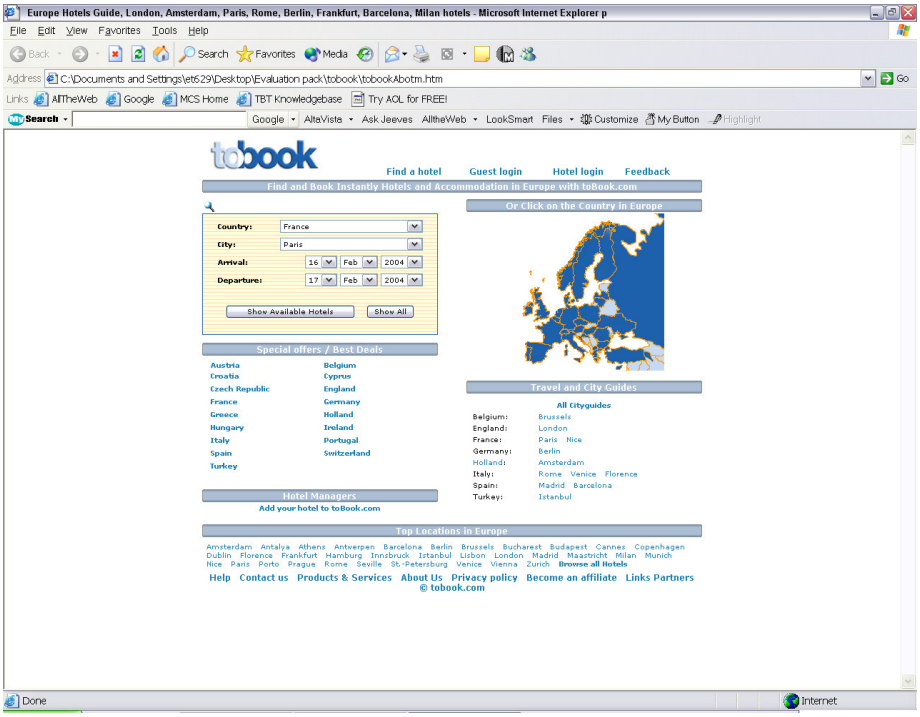


Figure 6-10 - Image of *Toobook* web page

The criteria for choosing homepages were that the pages should have:

- a design layout that fits within the computer screen (17 inches flat screen with a resolution of 1024 by 768 **pixels**) without requiring scrolling, and
- both a top- and bottom-page navigation bar where the 'About Us' link could appear.

The **pixel** size of the target design element 'About Us' along with its presentation format (emboldening) and position within navigation bar slightly varied across homepages. For example, the *Past Times* homepage had the three navigation bars at the bottom of the page and the 'About Us' link appeared in the middle one.

Eye movements were recorded using an ASL (Applied Science Technologies) 504 eye tracking remote pan-tilt camera (ASL, 2005) capturing eye movement data at a sample rate of 50Hz. The presentation of the stimuli was controlled by means of the GazeTracker™ software and presented on the screen viewed by participants from a distance of 55 cm from the screen. A detail description of the operation of the equipment is found in Chapter 4.

6.2.4 Procedure

The total duration of each participant's session was approximately 30 minutes. The session started by the researcher giving an introduction of the eye tracking equipment and the study to the participant. The participant completed a consent form and a background questionnaire. The participant was asked the following questions regarding the 'About Us' link: *Where do you look when you want to find information regarding the company? Where do you prefer to find it?* These

questions were aimed to collect information about the user's expectations and preferences regarding the placement of the 'About Us' link before the eye-tracking session.

The researcher then calibrated the eye tracking camera for the participant. The participant was asked to look at each web page and find the 'About Us' link. The participant was asked to say aloud where on the interface they found the 'About Us' link in order to indicate that the task had been completed, so that the researcher could press the 'enter' key on the keyboard for the next page to appear. Although there was no time limit for the task so as to encourage a natural navigation of the web page the participants were prompted to find the target link as quickly as possible. To avoid the researcher's reaction times influencing the data, the eye movement data were used as a measure of the task completion times. After the eye tracking session, the participants were asked: '*Where would you like to find the 'About Us' link? What do you think about the web pages you just saw? Was it easy to find the 'About Us' link'?* These questions aimed to collect information about the user's perceptions and preferences regarding the 'About Us' link after the repeated exposures session.

6.3 Data analysis

Both quantitative and qualitative data were gathered.

Data from the background questionnaire regarding the users' profile were grouped according to the participants' responses to structured questions whereas the data from the pre and post session questions were grouped according to the responses to semi-structured questions.

For the eye movement data the following metrics were applied for the data analysis:

- *Time to target fixation* which indicates the amount of time needed for the participant to fixate on the target link ('About us' link),
- *Location of initial gaze* which indicates where the user expect to find the target link,
- *Location of entry point* which indicates where the user looked first and,
- *Scan path reading* (sequence of fixations) which demonstrates the users' scan patterns.

In order to consolidate the data, a data organisation template was developed for each participant where all this information could be displayed. For example, see Appendix B. This was an effective way of identifying patterns across exposures and across participants.

6.4 Results

Repeated Exposures Effect

The study's first question aimed to examine whether users adapt their expectations of where to look when presented with repeated exposures of the 'About Us' link at the bottom of the page (an unconventional position). The descriptive statistics of the scores to target fixation across repeated exposures are shown in Table 6-3. There is a difference between the sum of time to target fixation for the first of the repeated exposures (Exposure 2) *179.10* and the last of the repeated exposures (Exposure 7) *38.7* in the predicted direction.

Exposures	N	Minimum	Maximum	Sum	Mean	Std.Deviation
Exposure 2	10	3.52	38.86	***179.10***	17.90	11.41
Exposure 3	10	1.44	44.97	120.16	12.01	13.11
Exposure 4	10	0.72	17.98	64.63	6.46	5.65
Exposure 5	10	1.53	31.64	83.26	8.32	9.99
Exposure 6	10	1.20	41.13	122.62	12.26	14.09
Exposure 7	10	0.91	6.80	***38.70***	3.86	1.61
Exposure 8	10	2.53	16.16	58.48	5.84	3.99
Exposure 9	10	1.80	22.86	91	9.09	6.79

Table 6-3 - Descriptive statistics of sums of time to target fixation for repeated exposures

A **non-parametric Trend test** (Page, 1963) was applied to test for a trend of learning where to look when presented with a sequence of six repeated exposures of homepages where the 'About Us' link appears on the bottom of the page.

A Page's L **trend test** (Page, 1963) on the ranked scores of time to target fixation for the repeated exposures revealed a significant trend across exposures: L (10, 6) =792, p<0.05. Time to target fixation decreases as the number of exposures increases is shown in Figure 6-11.

The eye tracking measures were complemented by the qualitative data as retrieved from the pre- and post–session questions. When the participants were asked before the eye tracking session where they expected to find the 'About Us' link they answered 'on top of the page' or it does not matter as long as they can see it. But, when the participants were asked after the eye tracking session where would they like to find the 'About Us' link they answered 'on the bottom of the page'. In addition to the eye tracking data where a trend of adaptation was found as exposures increased, the modification in answers from the pre- and post-session questions suggests an influence of change in preferences of where the 'About Us' link is expected to appear.

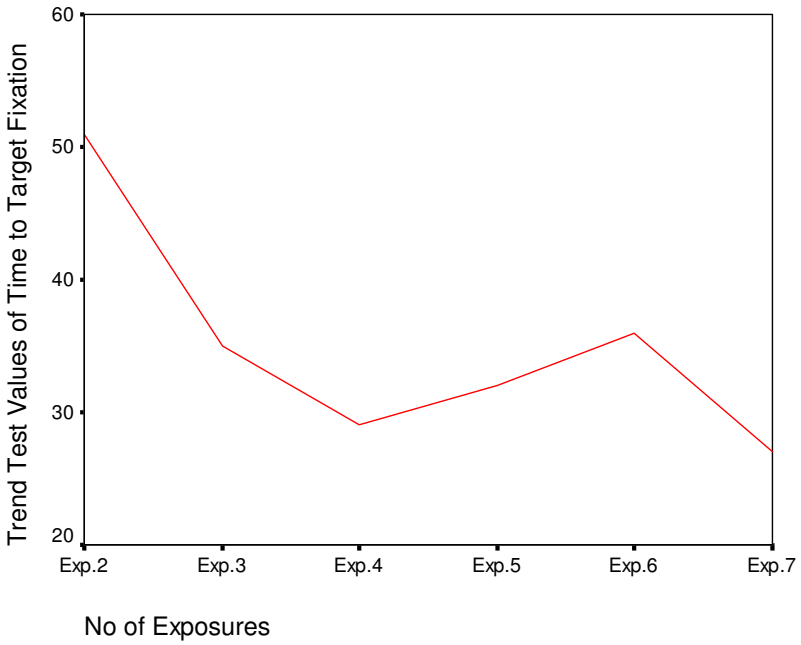


Figure 6-11 - Trend Test of time to target fixation across exposures

Alternative Design Layout Effect

The study's second question aimed to examine if users adapt quickly to alternative design layouts such as finding the 'About Us' link in an unconventional position.

A two-tailed paired t-test was used to establish whether there were significant differences between the time to target fixation before the repeated exposures ('About Us' link at the bottom of the page-Exposure 1) and after repeated exposures ('About Us' link at top of the page-Exposure 2) at the 5 per cent alpha level of confidence. There is evidence that the users found the target link quicker when the 'About Us' link was placed at the top of the page even after having seen it on the bottom of the page in the repeated exposures session as $t(9) = 3.35$, $p < .05$.

After transforming the raw scores using a two related samples Wilcoxon test, the distributions between the time to target fixation after the repeated exposures ('About Us' link at top of the page-Exposure 8) and after the alternative exposure ('About Us' link at bottom of the page again-Exposure 9) were compared at the 5 per cent alpha level of confidence. Despite the significant difference found when the 'About Us' link is presented on the top of the page, when it is then presented again at the bottom of the page there is a significant difference as $t(9) = 1.98$, $p < .05$.

The results indicate that when the 'About Us' link is placed on the top of the page, users find it quicker than when it is placed at the bottom of the page. This might be the effect of their previous experiences, as indicated in their self-reports and as could be expected based on the normal conventions of the 'About Us' link's placement that was found in the survey of leading e-commerce sites. But when presented with the 'About Us' link at the bottom of the page again after the repeated exposure a second expectation has been developed. Possibly this is caused by the consistent placement of the 'About Us' link on the bottom of the page over the repeated exposures.

Before and After Repeated Exposures Effect

Granka *et al*. (2004) used a location grid to analyse eye movements in relation to specific position on the user interface. This enabled the classification of initial expectations and visual attraction to specific design elements. Similarly Study 2 reported here used a location grid of six equal areas (see Figure 6-12) to determine the locations of 'initial gaze' and 'entry points'.

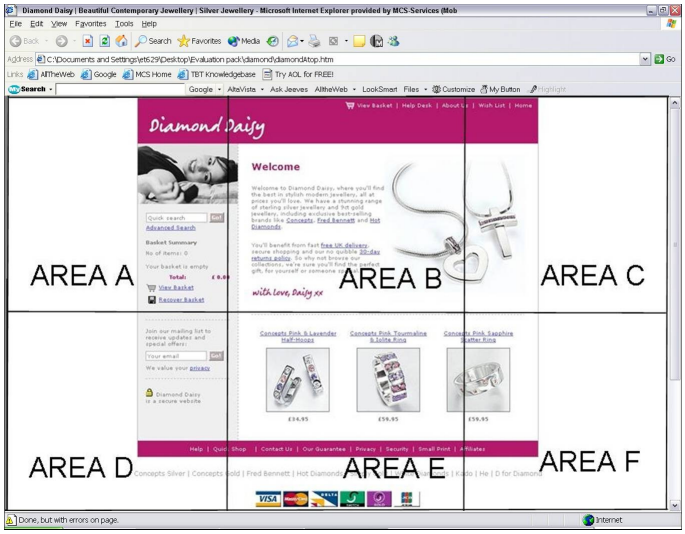


Figure 6-12 - Sample location grid to assess effect of design element location.

The location of 'initial gaze' for each participant was measured during the first 50 milliseconds of the homepage appearing on the screen to determine where the participant expected to find the 'About Us' link. 'Initial gaze' for all participants was always in areas A or B (that is the top left or top middle of the screen) for both exposures 1 and 10. None of the 'initial gazes' focused on the right side or bottom of the screen (areas C, D, E, F in Figure 2). This is more of an indication of similar visual search behaviour of initial gazes starting from the upper part of the page rather than any indication of user-adaptation across exposures. The very small amount of time (50 milliseconds) during which 'initial gazes' where measured might not have allowed the observation of any possible scan path modification. Therefore, the location of the 'entry point' for each participant was measured during the first 250 milliseconds of the homepage appearing on the screen to indicate where the participant first fixated. The 'entry points' were not consistent across participants and varied from homepage to homepage. Nevertheless none of the 'entry points' were in the right side of the screen (areas C and F in Figure 2).

This might be influenced by the visual attraction of specific design elements rather than purely consistent placement of the 'About Us' link.

By comparing the users' scan paths in Exposure 1 and Exposure 10, five users where found to have modified their scan patterns from the first to the last exposure after being presented with the repeated exposures session. For the first exposure they started their scan paths in the upper area of the screen (areas A and B) whereas after the repeated exposures session they started their scan paths in the lower part of the screen (areas D and E) suggesting user-adaptation after finding the 'About Us' link at the bottom of the page. When looking at the profile of these five users we found that they were frequent internet users (regularly throughout the day) and used the internet as part of their everyday work activities suggesting that they were highly skilled internet users which might explain their quick adaptation to consistent placement of specific design elements such as the 'About Us' link. The other five users were 'regular internet' users using the internet 2-3 times per day. Despite the trend of user-adaptation during the repeated exposures session, the variance of scores of time to target fixation lead to the examination of the effect of specific design characteristics of the target link for example, the high score for Exposure 6.

Effect of Design Characteristics

It is apparent that certain web pages required more time for the 'About Us' link to be identified. This was measured for each homepage, regardless of its place in the exposure order (Table 6-4). The *Past Times* homepage had the highest scores for the time to target fixation *190.22*, whereas the *Train line* homepage *27.23* had the lowest scores for the time to target fixation.

Web Pages	N	Minimum	Maximum	Sum	Mean	Std. Deviation
Cover4students	10	1.20	17.78	56.30	7.03	4.94
Hatton	10	1.88	32.80	99.19	12.39	12.00
Past-times	10	2.95	44.97	***190.22***	23.77	17.50
To book	10	2.23	22.86	56.45	7.05	7.25
Travel bag	10	1.53	13.94	46.77	5.84	4.15
Diamond	10	3.49	22.86	84.77	10.59	6.30
Hotel net	10	1.80	17.86	49.80	6.22	5.99
Saga	10	0 .91	17.98	54.57	6.82	6.30
Train line	10	1.44	5.21	***27.23***	3.40	1.13
Travelodge	10	0 .72	26.38	88.34	11.04	8.88

Table 6-4 -Descriptive statistics of sums of time to target fixation for homepages regardless exposure

When considering the possible effect of specific visual characteristics of design elements, three characteristics were explored:

- **pixel** size,
- text appearing bold or not, and
- the position of the 'About Us' link within the navigation bar.

The **pixel** size of the target design element 'About Us' link was measured using JRuler Pro (2004). No relationship between time to target fixation and **pixel** size was found ($r = -.189$, $p > 0.05$). So, small size in **pixels** did not always lead to increased time to target fixation (Figure 6-13).

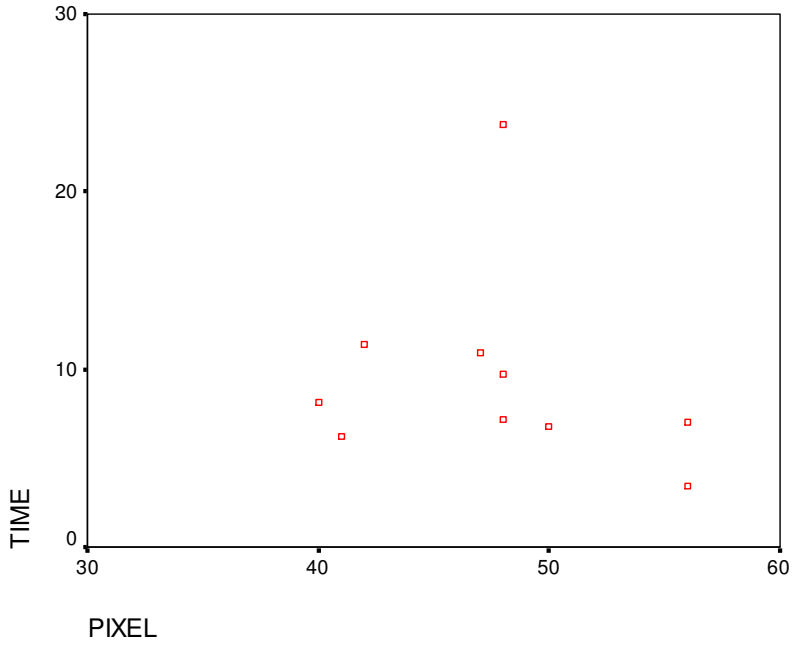


Figure 6-13 - Scatter plot of correlation between time to target fixation and pixel size

When the 'About Us' link did not appear in bold or distinctive colour (the formatting effects used in the example websites), the time to target fixation increased, implying that without distinctive formatting the target was more difficult to locate. The position of the 'About Us' link within the navigation bar did not appear to influence the time to target fixation. However, when more than one navigation bar appeared on the bottom of the screen, the time to target fixation increased.

6.5 Discussion

User-Adaptation

When placing the target link in a consistent position over a series of exposures, the results show that users adapt to consistent placement of the target link and improve their visual search performance. A trend was found of more exposures leading to decreased time to target fixation. Eye movement data were complemented by self reports of change in expectations and preferences of where the target link was expected to be found. The reported results comply with previous research as conducted by Ehret (2002, 2000) suggesting that users learn the locations of design elements over series of repeated exposures. On the contrary, McCarthy *et al*. (2003) had found no evidence that performance rates improve when the target link is placed in expected positions. They found that users adapt quickly to unexpected design layouts.

When placing the 'About Us' link at the top of the page as per the **norm** on most websites, as shown by the survey conducted as part of the reported study, the results indicate that users' visual search performance is quicker than when placing the target link at the bottom of the page. It is assumed that primacy effects of the users' exposure to web pages might have influenced their visual search performance, whereas influence of the repeated exposures develops as a secondary effect which was found to improve visual performance but does not override the effect from previous experiences.

Impact of Design Layout Characteristics

Distinctive formatting of design elements (for example emboldening and positioning within the navigation bar) also influences users' visual search behaviour

and shortens time to target fixation. However, this influence is secondary to the learning effects: the learning effects associated with consistency of placement of the design element over repeated exposures were observed regardless of the complexity and visual characteristics of the design elements.

Usefulness of Results for Designers and Researchers

This study should encourage designers to aim for consistency not only *within* a website but also consistency *across* websites. However, just placing design elements in similar positions on the user interface within or across a website does not guarantee efficient visual search. The presentation format of the design element influences user visual search behaviour as well. The results of the reported study indicated that design elements such as the 'About Us' link are not only located more quickly when presented in a consistent position over exposures, but also when they are presented in a distinctive format. For researchers the study demonstrates the usefulness of applying eye tracking as a usability evaluation technique for user interface designs. Although the reported study is limited to visual search of homepages, the methodology can be applied to other web pages and to a wide range of tasks.

Study 2 explored a broad usability concept and demonstrated the usefulness of eye tracking as a detector of detailed user behaviour. The results of Study 2 refine the concept and suggest specific recommendations for designers. The third and forth study of this research explore the influence of specific factors on visual search behaviour and performance. These factors are key factors of e-commerce environments such as the effect of text and images on visual search (Study 3) and the presentation format of key icons, such as the shopping cart (Study 4).

Chapter 7 Study 3 - Text-Image based Study

The soul never thinks without a mental picture.
Aristotle (384-322 BC)

The opportunity for Study 3 lies within the same reasoning of Studies 1 and 2; which is to test existing web design guidelines by conducting an eye tracking study. Study 3 focuses on key aspects of e-commerce sites. The text-image based study examines whether the number of images appearing on e-travel web pages influence users' visual search performance and behaviour. Moreover this study explores potential differences across text-based and image-based e-travel web pages in terms of visual search performance and behaviour.

7.1 Rationale of Study

HCI experts have argued whether image-based information is more effective than information presented in textual format (De Sanctis, 1984, Tullis, 1981). In contrast to De Sanctis and Tullis, Nielsen (2000) and Krug (2000) argue that images and graphics should be used as little as possible as they increase download times and distract users when completing their tasks.

While common opinion suggests that images are more effective for the comprehension of visual characteristics on the user interface, not all of the numerous studies conducted in this field have demonstrated images' usefulness over *text-based* information. For example, Williams (1999) tested 36 pilots in a flight simulator to assess their ability to decide which of two airports was furthest from a storm front, based on the manner in which information was presented on a

navigational display. The results supported the superiority of image-based information over text-based information display. Pilots were faster using the map display than using either of the two text displays. Nawrocki (1972) found no significant advantage to image-based information over *text-based* when participants were asked to remember previously presented information. Tullis (1981) suggested that the main conclusion to be drawn from a review of studies by Booher (1975) and Rigney and Lutz (1976) is not surprisingly that the effects of images on human performance are highly dependent on the task. Therefore, it is useful to ask what task factors influence this performance.

Nielsen and Tahir (2002) suggest allocating between five per cent and 15 per cent of the homepage **pixels** to images. A homepage allocating more than 15 per cent of the **pixels** to images is considered to be a bad design, for which users perceive images as distracting.

So, does the proportion of images appearing on web pages influence user performance? The study reported here used eye tracking to investigate how users search for target design elements visually across three different conditions:

- *text-based* (no images – just text-based information),
- *image low* (five per cent - 15 per cent of **pixels** allocated to images, as per design guidelines),
- *image high* (16 per cent - 45 per cent of **pixels** allocated to images – three times more than the suggested guidelines).

Informed by the previous studies, the work reported here aimed to address not only the interplay between text and images but the effect of the proportion of

images appearing on web pages in a specific sector such as e-travel. E-travel was chosen as the domain of interest, because:

- It is commonplace, with numerous companies and high profile on the web, and with broad usage across the experience spectrum;
- It is diverse, offering a variety of both services and 'products' for sale in many configurations; and
- It does not have a consistent style.

For example, users looking at e-tail websites expect and want to look at images of the products, whereas users booking a flight have no expectation of finding images, although users booking resort accommodation might well expect images.

A review of ten well-known UK e-travel sites found no consistency in the proportions of images and text presented on their homepages. Some of the websites included images on the homepage, whereas other websites included images further into the website.

This study aimed to revisit existing design guidelines (Nielsen and Tahir, 2002) that suggest limiting images to an optimum homepage screen share of only five per cent-15 per cent, by providing detailed empirical evidence of users' visual search behaviours. The objectives of the study are to compare visual search behaviour across different design presentation formats of web pages and across different tasks.

7.2 Method

The between-participants study using eye tracking collected four types of data:

- *A background questionnaire* to elicit the users' internet experience and typical usage,
- *Pre-session questions* to elicit users' expectations about the use of images and text on travel web pages,
- *Eye tracking data* to measure visual search behavior and performance, and
- *Post-session questions* to address perceptions and preferences about the use of images and text on travel web pages.

7.2.1 Study design

The between-participants design used three groups of ten participants each, or 30 participants in total (Table 7-1). Each group was assigned randomly to one of the three conditions:

- *'Text-based'* homepages where no images appear,
- *'Image-low'* homepages where five per cent-15 per cent of **pixels** are allocated to images on the homepage (suggested as optimum by Nielsen and Tahir, 2002),
- *'Images-high'* homepages where 16 per cent-45 per cent of **pixels** (three times more) are allocated to images on the homepage (suggested as bad design by Nielsen and Tahir, 2002).

Three commercial homepages were each modified to conform to each of the three conditions. Which homepage appeared in which task for any given user was assigned randomly.

The study was designed to address one of the issues raised in the literature: that task context influences visual search performance and behaviour. Study 3 compared performance and behaviour in three authentic tasks, typical of activities on e-travel sites:

- '*Please look at the web page and try to figure out what the site is about as quickly as possible.*'

This task was intended to prompt the users to scan the homepage relatively freely, in order to observe what design elements caught their attention.

- '*Please find the telephone number of the company and quickly tell me where you can see it on the screen'.*

The *rationale* for this task was to prompt the users to find a specific textual link, in order to compare scan paths and time to target fixation across the different conditions.

- '*Please find an offer for Paris and quickly tell me where you can see it on the screen'.*

The *rationale* for this task was to prompt the users to find a target link represented in the *text base,* image *low* and *image high* conditions by both textual and image-based elements, in order to compare scan paths and time to target fixation across the different conditions.

The study addressed the following questions:

- Q1. Does the proportion of images on a web page influence users' visual search?

- Q2. How does the absence/ presence of images on a web page influence users' visual search?

- Q3. How does the nature of the task influence users' visual search?

	Participants	Task 1	Task 2	Task 3	Condition
GROUP A	**10**	**Homepage1a, Homepage2a, Homepage3a appearing in a random sequence for each participant within each group**			**Text**
GROUP B	**10**	**Homepage1b, Homepage2b, Homepage3b appearing in a random sequence for each participant within each group**			**Image Low**
GROUP C	**10**	**Homepage1c, Homepage2c, Homepage3c appearing in a random sequence for each participant within each group**			**Image High**

Table 7-1 - Design for *image/text-based* Study

7.2.2 Participants

Thirty volunteers from the staff and postgraduate student population of the Open University participated in this study. Each participant received an honorarium of £10 to acknowledge their time spent in the study. The participants were divided evenly between the three conditions (*text-based*, *image low*, *image high)*; that is, each was assigned randomly to a group of ten for one of the conditions. The group for the *text-based* condition consisted of seven males and three females of ages ranging between 23 and 38. The group for the *image low* condition consisted of four males and six females of ages ranging between 23 and 45. The group for the *image high* condition consisted of four males and six females of ages ranging

between 22 and 56. All participants of all three groups used the internet either regularly throughout the day or every day. A detailed presentation of the participants' characteristics and their frequency and purpose of internet use can be found in Appendix B and are summarised in Table 7-2.

Sample Size	Characteristics	Text-based condition	Image low condition	Image high condition
Total number of participants 30 – 10 in each condition	***Age range***	23 - 38	23 - 45	22 - 56
	Sex	7 males – 3 females	4 males – 6 females	4 males – 6 females
	Frequency of internet use	3 every day 7 regularly throughout the day	3 every day 7 regularly throughout the day	1 every day 9 regularly throughout the day
	Purpose of internet use	To receive and write e-mail To buy or sell general products online To buy gifts to be delivered to friends or family To buy specialist equipment As an information source As a business medium To chat To listen to music To study online To read newspapers		

Table 7-2 -Characteristics of participants as elicited from background questionnaires

7.2.3 Stimuli and equipment

Three homepages were selected from the e-travel domain. They were selected to ensure that they:

- were not leading e-travel websites, in order to decrease the probability that the participants would have seen them before;

- shared a similar overall design layout, so that layout was not confounded with proportion of images; and

- could be modified in a way that would enable design elements such as images to be aligned.

The sites selected were:

- *First Call Travel* www.firstcalltravel.com (online travel agency),

- *Major travel* www.majortravel.co.uk (online travel agency), and

- *Need a hotel* www.needahotel.com (online accommodation booking).

Images of the selected web pages are shown in Table 7-3 below.

The overall proportion of images appearing on each screen was controlled for each of the conditions. The **pixel** size of the individual images appearing on the web pages was 72 x 65. See Appendix B for large screen shots of the modified web pages. Table 7-3 presents smaller screen shots of the stimuli.

Eye movements were recorded using an ASL (Applied Science Technologies) 504 eye tracking device. Detailed description of the apparatus is found in Chapter 4. Large images of stimuli screen shots could be found in Appendix B.

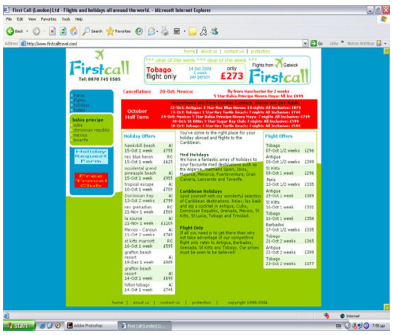
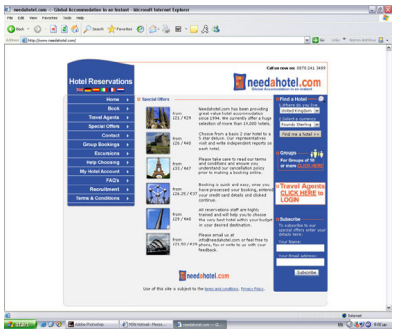
	www.firstcalltravel.com	www.majortravel.co.uk	www.needahotel.com
Text			
Image low			
Image high			

Table 7-3-Screen shots of stimuli web pages per condition

7.2.4 Procedure

The duration of one session including the briefing and **calibration** process was approximately 30 minutes. The session started with a spoken introduction to the participant of the eye tracking equipment and the study. The participant completed a consent form and a background questionnaire which captured age, gender,

previous internet experience, and frequency and purpose of internet use. The participant was asked the following scripted questions regarding the presence or absence of images on e-travel websites:

- *Do you buy tickets or book hotels online?*
- *What kinds of sites do you like and why?*
- *What happens when you have images on the homepages of travel websites?*
- *Do you think it helps?*
- *Do you like it?*
- *How about having no images?*

These questions aimed to collect information about the user's expectations and preferences regarding the presence or absence of images on e-travel websites.

The researcher calibrated the eye tracking camera for the individual. The participants were asked to look at three web pages and complete one task for each. The presentation order of the web pages was randomised, but the task order was the same for all participants.

After the eye tracking session, the participants were asked questions about web pages they had seen during the eye tracking session. The scripted questions were:

- *First of all how did you feel about the tasks?*
- *Was it difficult to complete any of the tasks?*
- *Do you have any preference of the three web pages you just saw?*

- *Any reason for that?*

- *Here I have print outs of more web pages and the ones you just saw. Have a look and tell me what do you think about them?*

- *Which ones do you prefer and why?*

- *Did you find it difficult to recognise the tower of Eiffel in the images?*

These questions aimed to collect information about the user's preferences regarding the presence or absence of images on e-travel websites after the eye tracking session, and to compare those expectations and preferences to those elicited by the pre-session questions.

7.3 Data analysis

The analysis of the data for the reported study focused on two aspects of visual search (Figure 7-1):

1. Statistical analysis of *performance,* as measured by time to target fixation:

1.1 Comparison by task: Compared performance on each task between the presentation format of the three conditions.

1.2 Comparison by condition: Compared performance within the presentation format of the three conditions.

2. Mixed quantitative and qualitative analysis of *behaviour,* as indicated by gaze duration within *look zones*, transition of fixations between design elements, and scan path patterns:

2.1 Comparison by task: Compared behaviour on each task between conditions

2.2 Comparison by condition: compared behaviour across tasks within each condition.

2.3 Comparison by participant: compared behaviour across participants.

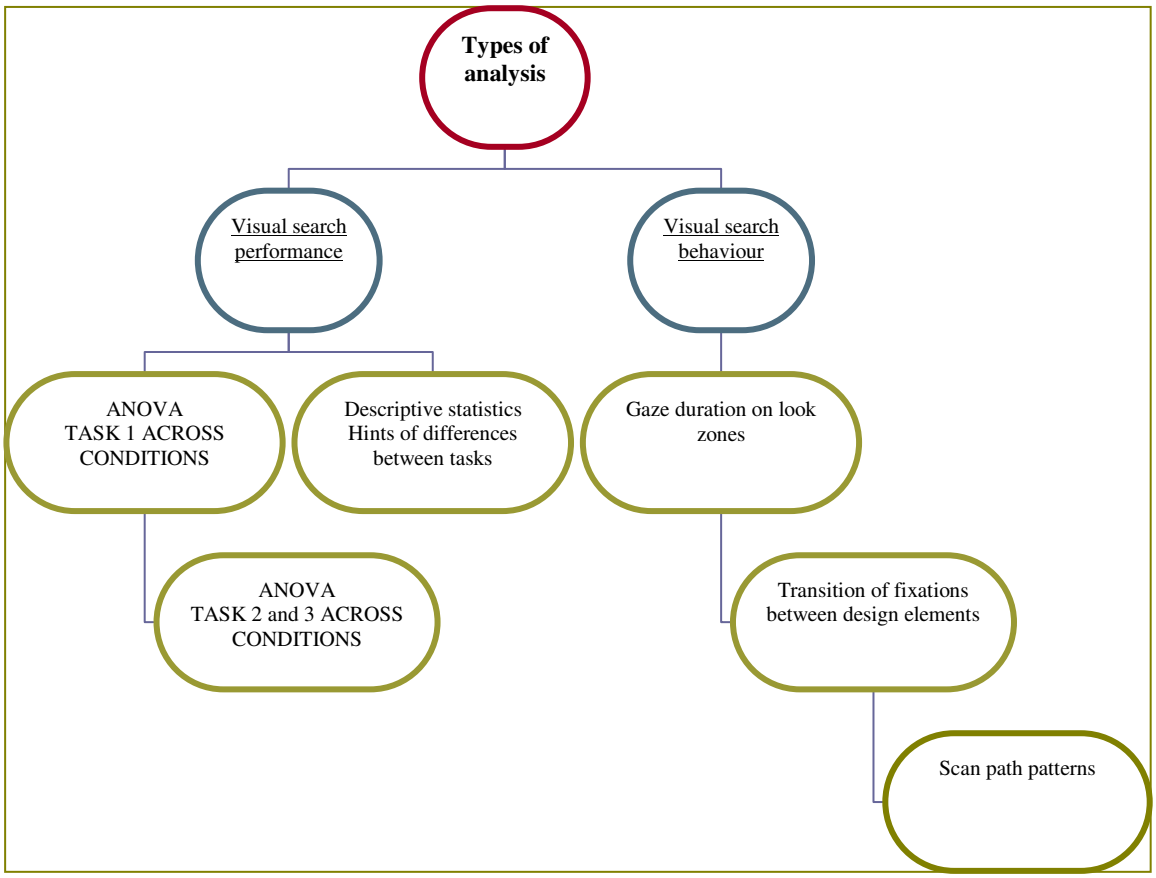


Figure 7-1 -Diagram of types of analysis

7.4 Results

The study's questions aimed to examine whether the users' visual search is influenced by two main factors:

- the design layout and
- the task itself

When talking about visual search the focus is on two main concepts:

- visual search performance, and
- visual search behaviour.

The results are presented per concept, followed by a brief interpretation and discussion.

Task 2 – Telephone Number and Task 3 – Offer for Paris

'Please find the telephone number of the company and quickly tell me where you can see it on the screen.'

'*Please find an offer for Paris and quickly tell me where you can see it on the screen.'*

The descriptive statistics of the scores to target fixation across the three conditions for Task 1 and Task 2 are shown in Table 7-3.

Presentation Layout Factor		N	Mean	Std. Deviation
TEXT	Task 2	10	7.07	4.09
	Task 3	10	6.86	3.30
IMAGE LOW	Task 2	10	8.44	7.05
	Task 3	10	6.81	2.91
IMAGE HIGH	Task 2	10	9.04	6.22
	Task3	10	5.11	1.01

Table 7-3 - Descriptive statistics of sums of time to target fixation (in milliseconds) across three conditions for Task 2 and Task 3

A three-by-two analysis of variance (ANOVA) was calculated on participant's scores of time to target fixation across the three presentation layout factors. The analysis showed an interaction effect but no significant difference was found F (2, 17) = .011, $p > .05$. Figure 7-2 illustrates the interaction effect between the two tasks and the presentation layout of the three conditions.

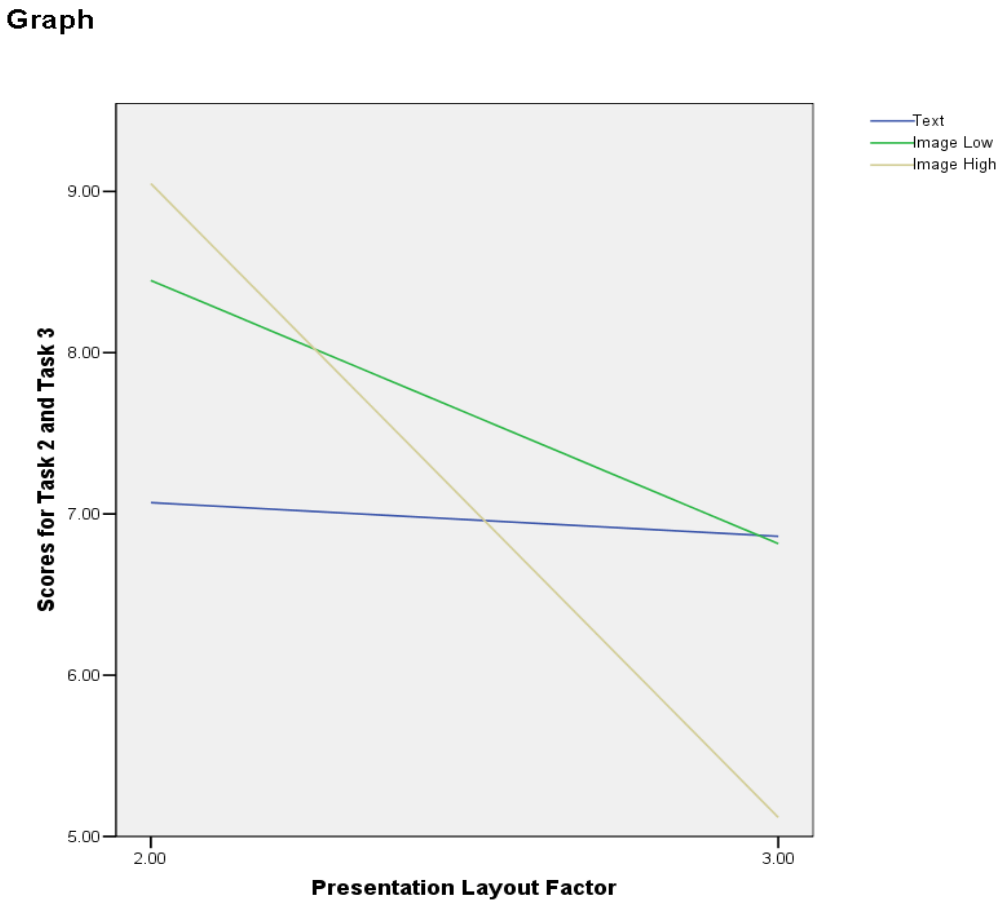


Figure 7-2 - Mean of time to target fixation for Task 2 and 3 across three presentation layout conditions

As Task 1 was different in nature from Task 2 and Task 3, it was not included in the analysis of variance with the other two tasks. It was calculated in a one-way analysis of variance as discussed below.

Task 1 – The purpose of the site

'Please have a good look at the web page and quickly tell me what the site is about.'

The descriptive statistics of the times to task completion across the three conditions for the first task are shown in Table 7-4 and Figure 7-3.

Conditions	N	Minimum	Maximum	Sum	Mean	Std. Deviation
TEXT	10	8.70	27.78	147.75	14.77	6.57
IMAGE LOW	10	4.81	27.22	141.81	14.18	7.52
IMAGE HIGH	10	7.41	27.22	141.54	14.15	6.53
Total	30					

Table 7-4 - Descriptive statistics of sums of time to target fixation (in milliseconds) across three conditions for Task 1

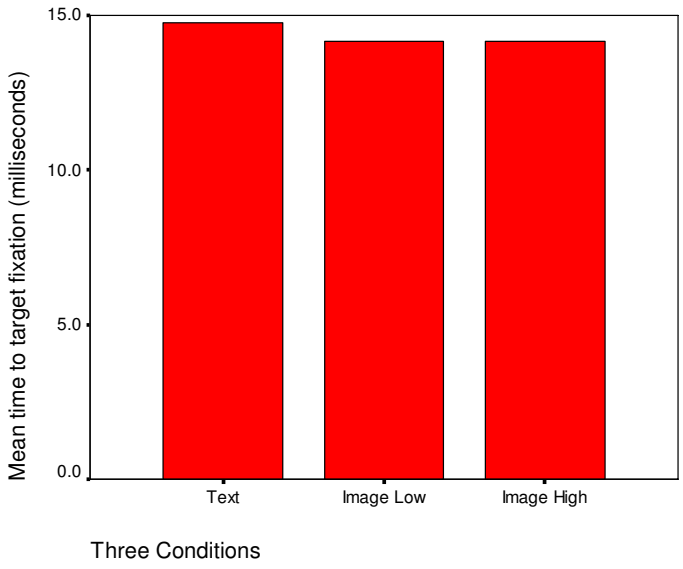


Figure 7-3 - Mean of time to target fixation for first task across three conditions for Task 1

A one–way analysis of variance (**ANOVA**) was calculated on participants' time to target fixation across the three conditions for the first task. The analysis showed that the small differences between conditions were *not* significant differences, *F* (2,

27) = .026, $p > .05$, indicating that the participants were able to complete the task in similar times in all three conditions.

Time to target fixation across tasks within conditions

The comparison of time to target fixation for each task within the different conditions did not indicate any significant differences. What of the comparison of time to target fixation for each task *across* the different conditions? Given the difference in the nature of the three tasks, a statistical comparison such as **ANOVA** was not appropriate. However, the descriptive statistics for time to target fixation for each task across conditions (Table 7-5) suggest differences between the first task (*14.36*) and the other two tasks (*8.18* and *6.26*). This is a matter for further investigation.

Tasks	N	Minimum	Maximum	Sum	Mean	Std. Deviation
TASK 1	3	14.15	14.77	43.10	14.36	.349
TASK 2	3	7.07	9.04	24.55	8.18	1.00
TASK3	3	5.11	6.86	18.78	6.26	.996
Total	9					

Table 7-5 - Descriptive statistics of sums of time to target fixation (in milliseconds) for each task across conditions

Discussion of performance results

There were no significant differences in performance between conditions. This is surprising, given the emphasis in the published guidelines on how images and text are used. The tasks were designed to require information presented as text only (Task 2) or as image only (Task 3, *image low* and *image high* conditions). One might have expected that a match between the type of information the task

required and the condition (that is, *text-based* and telephone number) would lead to a discernable advantage in performance.

Given the differences in the tasks, it is difficult to make a meaningful comparison of performance between them. However, a question was raised about the impact of the nature of the task on performance: it may be that exploratory, descriptive tasks (that is, Task 1) take longer than focussed fact-finding tasks (such as Tasks 2 and 3).

Scan path visual search behaviour

Three analyses were conducted to explore visual search patterns:

- Quantitative analysis of gaze duration in *look zones*
- Quantitative analysis of transition of fixations between *look zones*
- Qualitative analysis of scan path patterns

Gaze duration in *Look Zones*

The GazeTracker™ software allows the classification of *look zones*, areas of interest on the screen, such as navigation bars, collections of images, or promotional text. *Look zones* were customised to the stimuli, that is, they were defined individually for each web page. The categorization scheme for the types of *look zones* was a variant of the scheme used by Nielsen and Tahir (2002) to classify design elements on homepages: *Logo, Navigation, Promotional text, Content text, Image, Telephone number, Address,* and *Unused space.* An example of the classification of design elements is shown in Figure 7-4 and an example of the 3D data presentation of percentage of gaze duration are shown in Figure 7-5.

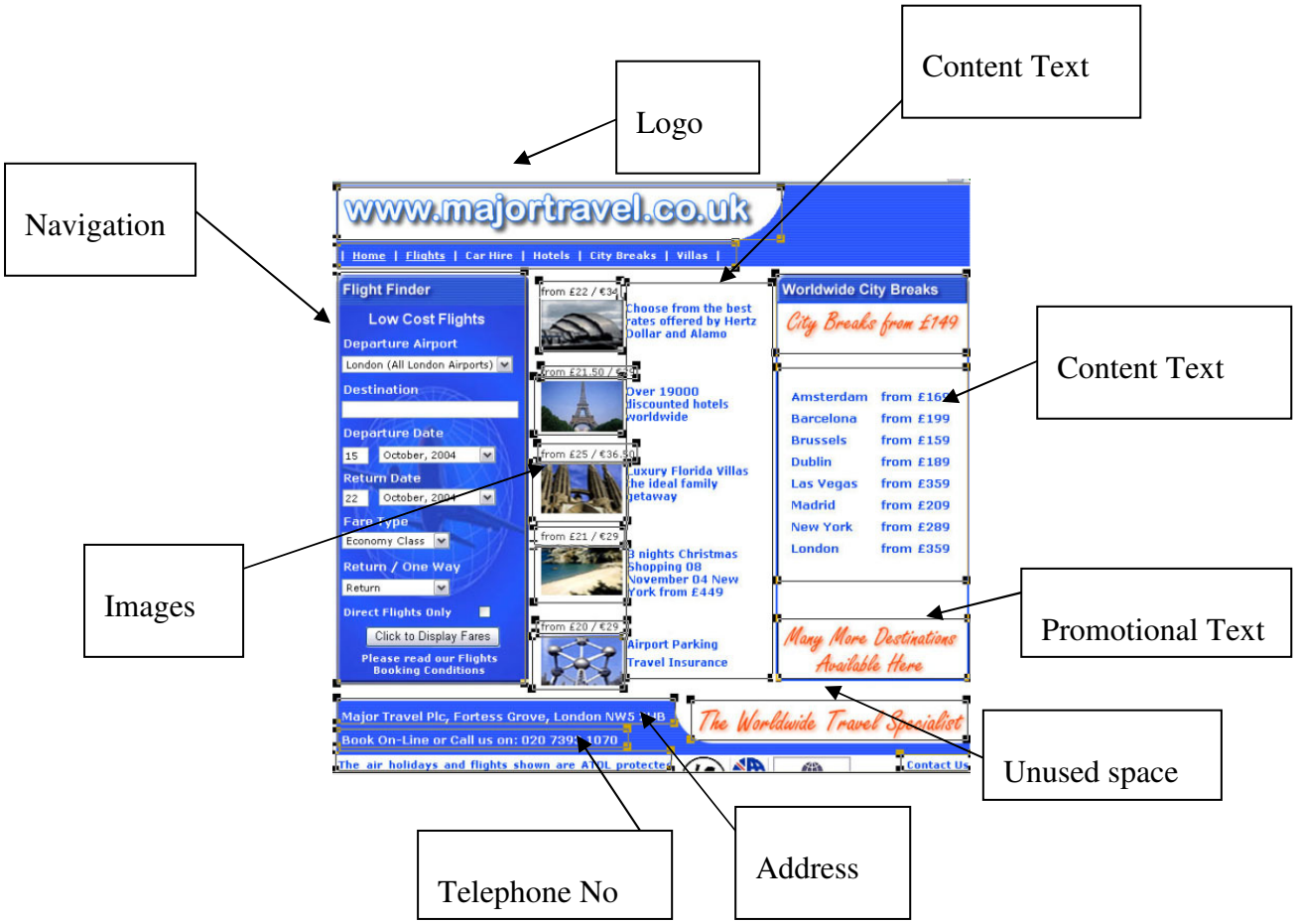


Figure 7-4 - An example of classification of types of *look zones*

Figure 7-5 - An example of classification of types of *look zones* and 3D data presentation of percentage of gaze duration within the defined *look zones* for one participant

According to Fitts, Jones and Milton (1950), frequency and duration of eye movements should be treated as separate metrics, with duration reflecting difficulty of information extraction, and frequency reflecting the importance of that area of the display. As the reported study was primarily interested in the difficulty of information extraction across the three different conditions, a duration measure was adopted. The basic unit of analysis was gaze duration which reflects the time spent looking at defined *look zones*. The table in Appendix B shows the percentage of time spent looking in *look zones* averaged across participants for each task per condition.

Transition of fixations between *look zones*

A conditional probability calculation was used to analyse the sequential dependencies between pairs of events. Conditional probability, often referred to as transitional probability, quantifies the extent to which one behaviour (the first or criterion event) is temporally related to another (the second or matched event) (Marion *et al*. 2003).

Goldberg and Kotval (1999) refer to the transition matrix as the frequency of eye movement transitions between defined *look zones*. Frequent transitions from one region of a display to another indicate inefficient scanning with extensive search.

The transitions matrix (presented in a table in Appendix B) shows a tabular representation of the percentage of most frequent pairs of fixations to and from each defined look zone as per the general classification of design elements given above. A high level classification of categories of design elements was used for the

transition fixation analysis. Similarly as used for the gaze duration analysis look zones were defined for the *Logo, Navigation, Promotional text, Content text, Image, Telephone number, Address,* and *Unused space.* For the transition fixation analysis three main categories were classified:

- 'text' which refers to *content text,*
- 'images' which refers to *images*,
- 'other' which refers to typographically-distinctive textual elements such as: *navigation, promotional text, logo, telephone number, address* and *unused space.*

The interest was to identify transitions of fixations from one design element to another in terms of their relevance. For example, fixations on design elements followed by fixations on the same, different and related design elements (Appendix B).

Scan path analysis of visual search behaviour

A qualitative analysis of the scan paths was conducted in order to classify possible visual search patterns. The classification of different types of scan paths emerged from the scan path data themselves. Four characteristic scan paths were identified:

Processing was defined quantitatively as ten or more successive fixations on one single design element, such as a navigation area (Figure 7-6).

Figure 7-6 - Representative scan path, classified as *processing*, superimposed on a stimulus web page in the *image low* condition

Anchoring was defined as successive fixations on different yet related design elements. For example, fixations on different images that are presented in a single column and appear one underneath the other Figure 7-7).

Figure 7-7 - Example scan path classified as *anchoring* superimposed on a stimulus web page in the *image high* condition

Random scanning was defined as single fixations on different and non-related design elements. For example, a fixation on text in area A followed on an unrelated image in area C (Figure 7-8).

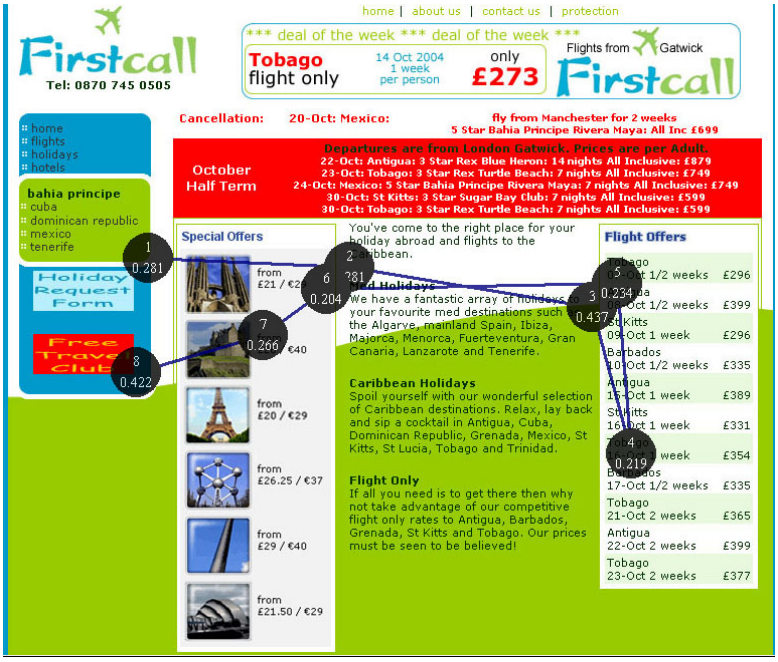


Figure 7-8 - Representative scan path classified as *random scanning* superimposed on a stimulus web page in the *image low* condition

Focused scanning was defined quantitatively as very small total numbers of fixations fewer than five fixations for a task. These fixations were often related to the target task (Figure 7-9).

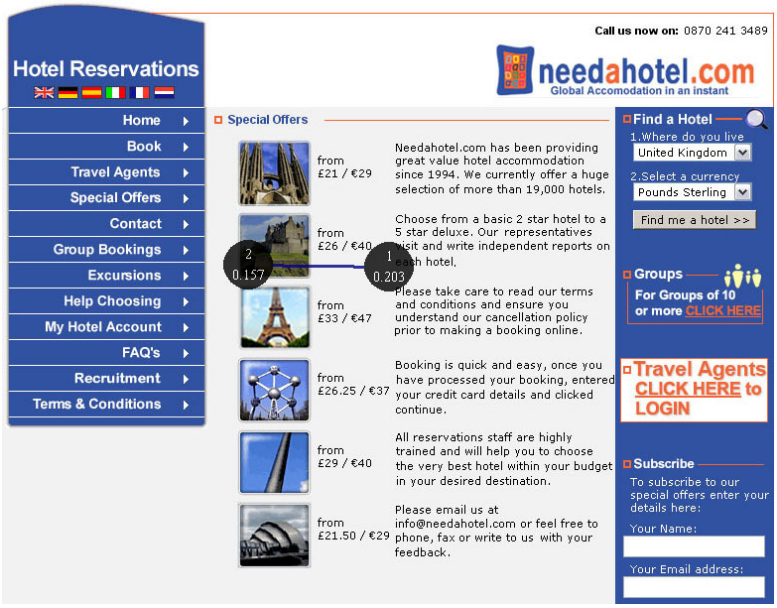


Figure 7-9 - Representative scan path classified as *focused scanning* superimposed on a stimulus web page in the *image low* condition

Observations from all three types of analysis

Detailed outcomes from each type of analysis can be found in tables in the appendix section: For gaze duration on *look zones* Appendix B, for transitions of fixations Appendix B, and for scan path patterns Appendix B. The analysis of visual search behaviour was best treated as a mosaic. Analysis of single relationships (such as the relationship of gaze duration to task) were not individually informative. However, the composite of the analyses enabled a better understanding of visual search patterns.

Most frequent and high scores have been reported in a combined format in Table 7-6, with outcomes from all three types of analysis. This composite table allows the following observations

- across conditions,

- across tasks, and

- across participants.

Outcomes from different types of eye movement data analysis				
Conditions	**Types of analysis**	**Task 1 – what is the site about**	**Task 2 – telephone No**	**Task 3 – offer for Paris**
Text-based	**Gaze duration**	More time was spent looking at the *look zones* which consisted of content text	More time was spent looking at the *look zones* which consisted of content text	More time was spent looking at the *look zones* which consisted of content text
	Transition of fixations	The most frequent transition was fixations on 'other' followed by fixations on same 'other'	The most frequent transition was fixations on 'other' followed by fixations on same 'other'	The most frequent transition was fixations on 'text' followed by fixations on same 'text'
	Scan path patterns	6 out of 9* participants performed a 'processing' type scan path pattern	4 out of 9* participants performed a 'processing' type scan path pattern	5 out of 9* participants performed a 'focused scanning' type of scan path
Image low based	**Gaze duration**	More time was spent looking at the *look zones* which consisted of content text	More time was spent looking at the *look zones* which consisted of promotional text	More time was spent looking at the *look zones* which consisted of navigation design elements
	Transition of fixations	The most frequent transition was fixations on 'other' followed by fixations on same 'other'	The most frequent transition was fixations on 'other' followed by fixations on same 'other'	The most frequent transition was fixations on 'text' followed by fixations on same' text'
	Scan path patterns	6 out of 10 participants performed a 'processing' type scan path pattern	3 out of 9* participants performed an 'anchoring' and 'focused scanning' type of scan path pattern	4 out of 9* participants performed an 'anchoring' type of scan path pattern
Image high based	**Gaze duration**	More time was spent looking at the *look zones* which consisted of images	More time was spent looking at the *look zones* which consisted of images	More time was spent looking at the *look zones* which consisted of images
	Transition of fixations	The most frequent transition was fixations on 'other' followed by fixations on same 'other'	The most frequent transition was fixations on 'other' followed by fixations on same 'other'	The most frequent transition was fixations on 'images' followed by fixations on different 'images'
	Scan path patterns	5 out of 9* participants performed a 'anchoring' type of scan path pattern	4 out of 9* participants performed a 'anchoring' type of scan path pattern	5 out of 9* participants performed a 'anchoring' type of scan path pattern

* There are a total of 9 participants in some cases because of missing data. This applies for all the cases in the table where frequencies refer to 9 participants

Table 7-6 - Outcomes of different types of analysis across conditions and tasks

Across conditions comparisons

When looking at the outcomes of each type of analysis for the three conditions it becomes evident that:

- The *look zones* that attracted high gaze duration varied between conditions.

During the *text-based* condition more time was spent looking at 'content text' (even though the target information for Tasks 2 and 3 were in other zones). During the *image low* condition the *look zones* which attracted high gaze duration varied for each task; high gaze duration associated with the *look zones* containing the target information. During the *image high* condition more time was spent looking at 'images' for all three tasks (again, even though the target information for Tasks 1 and 2 were in other zones).

This suggests that participants in the *image low* condition varied their visual behaviour to suit the task, whereas participants in the other conditions did not. This observation complies with outcomes of previous research (Antes, 1974, Mackwort and Morandi, 1967) which suggested that users fixate on regions of visual scenes that contain the dominant feature, so for example, in the *text-based* condition 'content text' (the *look zones* with high percentage of gaze duration) attracted longer gaze duration across all three tasks. Similarly in the *image high* condition, where the proportion of 'images' appearing on the screen was higher than the other two conditions (45 per cent) attracted longer duration across all three tasks (Table 7-6). The distribution of design elements in the *image low* condition was relatively even, so that none emerged as the dominant feature.

- In most cases, the most frequent fixation transitions were to the same design element.

Within Task 1 and 2, for all conditions, the most frequent transitions of fixation were to the same design element (from 'other' to the same 'other', where the so-called 'other' design elements refer to typographically-distinctive textual elements such as navigation bars, promotional text, logos.). The same was true for

Task 3 for the *text-based* and *image low* conditions (in this case, from 'content text' to same 'content text'). The exception to this pattern was the *image high* condition in Task 3, in which the most frequent transitions were from one 'image' to a different 'image'. This similarity in transitions of fixations across conditions indicates that users tend to look at design elements more than once before proceeding to another design element.

Previous research (Goldberg and Kotval, 1999) interpreted frequency of fixation transitions as an indicator of inefficient scanning, suggesting that the subject of repeated fixations is difficult to process. However, based on the behaviours reported here, it is conjectured that frequency of fixation transitions might indicate not *harder* or *less efficient* processing, but possibly just *more* processing – that is, the repeated fixations may indicate areas that contain goal-related information, which users perceive as useful. Therefore, the scanning is purposeful, rather than inefficient; the users appear to spend time 'reading' rather than just recognising. This conjecture is supported by the scan path analysis, which shows overall scanning behaviours that focus on information-rich areas relevant to the task. This applies across all tasks and all conditions, encompassing even the *image high* condition in Task 3, in which the 'images' contained relevant destination information. This conjecture and the evidence that supports it are consistent with the findings of Radach *et al*. (2003), who also used eye tracking on textual and pictorial advertisements. They found that the task has profound consequences for the distribution of viewing time and fixation positions. A greater proportion of viewing times was devoted to the elements that were more relevant to the task.

- Scan path patterns varied between all three conditions.

During the *text-based* and *image low* conditions most participants applied a 'processing' type of scan path (example in Figure 7-6) for Task 1, whereas in the *image high* condition most participants applied the 'anchoring' type of scan path (example in Figure 7-7). Likewise the scan path patterns varied across conditions for Task 2 and Task 3.

As discussed earlier, the high proportion of 'content text' in the *text-based* condition, and the high proportion of 'images' in the *image high* condition might be interpreted as dominant features that attract the gaze (following Antes, 1974, Mackwort and Morandi, 1967). From this, we might expect also to see 'anchoring' in these two conditions. However, 'anchoring' was mainly apparent in the *image low* and *image high* conditions (with only two examples in the *text-based* condition, in Task 1). One explanation lies in the interaction between visual strategies: between visual dominance and task relevance. It is conjectured that, although the gaze may be attracted initially to dominant features, repeated fixation relates to task relevance (as will be discussed further in the next section). If the dominant features do not evidently contain information relevant to the task – or if they are not sufficiently discriminable for their relevance to be apparent – then attention and strategies will shift.

Across tasks comparisons

In the *image low* condition, all three analyses show differences in behaviour associated with the different tasks. In the *text-based* and *image low* conditions, Tasks 1 and 2 shows similar behaviours, but behaviour differs in Task 3.

Yarbus (1967) found that a user's intentions, expectations and strategies for scanning a visual scene are moderated by the particular task and stimulus materials which are presented. As discussed above, the *image low* condition has a design layout which balances the proportions of different design elements, compared to the other two conditions which emphasise either 'content text' or 'images'. This may have had the effect of minimising a competing strategy: the effect of the task itself, and the variation of visual behaviour to match the task, becomes more evident without the influence of a 'dominant feature'.

Stenfors *et al*. (2003) suggest that the impact of dominant features may be mitigated by user experience. They report that users employ 'passive avoidance', ignoring design elements they do not consider relevant to their search. Highly salient visual cues such as high colour contrast and motion that usually attract attention appear to have no effect on the frequent internet user. The results reported here (particularly the consistency of behaviour with Tasks 1 and 2 for the *image low* and *image high* conditions) are somewhat at odds with Stenfors *et al*.'s findings, suggesting that experienced users may not be immune to dominant features. The alternative conjecture, offered above, is that although the gaze may linger initially on dominant features, attention (reflected by repeated fixation) moves effectively to task-relevant features. This is supported by the lack of significant performance differences between conditions. The impact of user experience is beyond the scope of this research.

Across participants comparisons

Eleven individuals out of 30 show a consistency of scan path patterns across tasks, suggesting a consistency of visual search behaviour across different tasks.

The similarity of scan path patterns is not echoed in the other measures (gaze duration and transitions). For example, the 16 who applied 'focused scanning' reported that they use the internet frequently throughout the day as part of their job. The users' internet experience in terms of frequency of use seems to be an influential factor in scan path patterns.

Pre- and post-session questions

Before and after the eye tracking session the participants were asked a set of open-ended questions regarding:

- previous experience using e-travel websites,
- expectations of having images on e-travel websites, and
- preferences or dislikes for images on websites.

The post-session interview included the presentation to the participants of print-outs of the web pages in all three conditions. The participants were asked to choose the one they liked the most. Table 7-7 and 7-8 summarises the responses across participants.

Conditions		Pre- session Responses				
		Previous experience using e-travel	**Examples of use of websites**	**Absence/presence of images**	**Images perceived as helpful**	**Preferences of images**
Text-based	**Group A**	10/10 booked tickets online 6/10 also booked accommodation online	8/10 Easy jet 7/10 Ryan air 2/10 Thompson 2/10 BA 2/10 National Express 2/10 Virgin 2/10 Fly B 2/10 Austrian airlines	7/10 have not seen images on travel websites 10/10 have seen images on accommodation sites 4/10 have seen images on promotional websites (i.e., Expedia, Opodo)	2/10 Yes for travel websites 4/10 Only for accommodation 2/10 Do not t mind 2/10 No	5/10 OK to have no images 4/10 Only for accommodation 1/10 Do not mind
Image low	**Group B**	10/10 booked tickets online 5/10 also booked accommodation	10/10 Ryan air 9/10 Easy jet 2/10 BA 1/10 BMI 1/10 Expedia 1/10 Google	5/10 have not seen images on travel websites 9/10 have seen images on accommodation sites 1/10 have seen images on promotional websites (i.e., Expedia, Opodo)	6/10 Only for accommodation 4/10 No	5/10 OK to have no images 3/10 Only for accommodation 2/10 Do not t mind
Image high	**Group C**	9/10 booked tickets online 5/10 also booked accommodation	8/10 Ryan air 7/10 easyjet 1/10 Ticket master 1/10 Canadian airlines 1/10 Euro star 1/10 National express 1/10 Cyprus airways 1/10 BA	5/10 have not seen images on travel websites 3/10 have seen images on accommodation sites 2/10 have seen images on promotional websites (i.e., Expedia, Opodo)	8/10 Only for accommodation 2/10 Yes 2/10 Do not mind	5/10 OK to have no images 4/10 Only for accommodation 1/10 Do not mind

Table 7-7 - Consolidated data pre-session questions presenting the frequency of responses to semi-structured questions across participants

Pre-session responses

All of the participants had previously booked flights online but not all had booked accommodation online. The most frequently visited e-travel web pages

were Easy Jet and Ryan Air. More participants reported having seen images on accommodation websites than on flight-booking websites. Similarly images were perceived as more helpful for accommodation, whereas half of the participants for each group did not care whether travel websites had images or not.

Post-session responses

Interestingly, those who said that they did not mind not having images on travel websites before the eye tracking session changed their preferences after the eye tracking session when shown print-outs of the stimuli web pages. They preferred pages with images. One participant said: '*I know I said I don't like images, but these here are just about right, not too many to distract me, I like them*'. Another participant said: '*I like the one with the images now, the images stand out, and the text is boring*'.

Six participants reported that Task 2 (find the telephone number) was confusing, because the location of the number on the stimuli sites did not match their expectations. One of the stimuli websites *Need a hotel* had the telephone number on the top right of the page. The participants said it was not straightforward to find it, as they expected it to be on the bottom of the page, where they had seen it previously on other websites.

First call and *Major travel* were selected as the most preferred web pages because the colour combinations and the design layout were perceived as more organised.

Conditions		Post- session Responses			
		Difficulty on completing tasks	**Preferences on stimuli web pages**	**Preferences on conditions**	**Difficulty with image of Eiffel tower**
Text-based	**Group A**	3/10 Confused when telephone number placed on the top right of screen 7/10 No difficulty	5/10 First call 4/10 Major travel 1/10 Need a hotel	3/10 Image low 1/10 Text 6/10 Image high	10/10 None
Image low	**Group B**	2/10 Confused when telephone number on the top right of screen 1/10 First call confusing name 7/10 No difficulty	4/10 First call 4/10 Major travel 2/10 Need a hotel	4/10 Image high 3/10 Image low 3/10 Text	10/10 None
Image high	**Group C**	1/10 First call confusing name 1/10 Confused when telephone number on the top right of screen 8/10 No difficulty	5/10 First call 4/10 Major travel 1/10 Need a hotel	4/10 Image high 4/10 Image low 2/10 Text	10/10 None

Table 7-8 - **Consolidated data post-session questions presenting the frequency of responses to semi-structured questions across participants**

The data from the background questionnaire and pre- and post-session questions have enabled a better understanding of the relationship between the users' preferences and visual search behaviour. Despite users having strong preferences of travel websites without images and only perceiving images as useful for accommodation websites, they actually looked at the images during the eye tracking session.

Eye movement data can provide easily comparable quantitative metrics for objective design interaction, but in combination with qualitative data that provide information about the user's preferences and expectations, a more accurate 'picture' is given of the user's visual search performance and behaviour.

7.5 Discussion

The opportunity for the reported study lies within the broader area of web design guidelines for e-commerce sites. Existing web design guidelines have been derived mostly from expert inspections rather than evaluation involving the users themselves (Ivory *et al*. 2001).

A specific guideline from Nielsen and Tahir (2002) reports that the proportion of **pixels** in images of a screen should not exceed 15 per cent in order to allow optimum user performance.

The reported study examined the effect of:

- The proportion of images appearing,
- The absence/presence of images on web pages, and
- Tasks performed

on web pages of e-travel sites.

The primary evaluation technique was eye tracking and we mainly focused on two aspects of visual search: *performance* and *behaviour.*

Visual search performance

On the measure of visual search *performance* it was found that the image low condition (as per guideline by Nielsen and Tahir, 2002) did not lead to faster search times. No significant differences were found in any of the three conditions, which lead to the conclusion that the proportion of images and the absence/presence of images on web pages does not lead to quicker performance.

Visual search behaviour

The measure of visual search *behaviour* gives an extra dimension to the results and beyond performance. While no differences in task times were found across conditions, differences were found between tasks, in terms of visual search behaviour, in only one condition: image low (as per guideline by Nielsen and Tahir, 2002). Dominant features such as 'text' and 'images' in the other two conditions *text-based* and *image high* were found to be a stronger influential factor.

The key conjecture made from the results is that despite not finding performance differences by condition nor by task, behaviour is orientated to both condition and task. Users apply interaction strategies that are influenced both by dominant features and task relevance.

Hints leading to further work

Overall, the results of the reported study provide a broad support of the design guideline suggested by Nielsen and Tahir (2002), but based on interaction strategies found in behaviour rather than performance. The *image low* condition minimised the impact of the dominant feature and enabled task related strategies. This raises a question about the discriminability of design elements such as the size or placement of images.

Further work might focus on other possible factors that influence visual search behaviour. A study which might control for colour or placement of images might provide useful insights of how discriminability of design elements such as images can be an influential factor in visual search.

Chapter 8 Study 4 - Shopping cart Study

Men are disturbed not by things that happen, but by their opinion of things that happen.
Epictetus (55 - 135 CE)

The previous chapter described a study in which the amount and presence/ absence of images appearing on homepages of e-travel websites varied, whereas this chapter presents a study where the presentation format of a target link, such as a shopping cart, has been modified in three different presentation formats.

8.1 Rationale of Study

The presentation format of navigation links varies across websites but the most popular are based on the basic formats of: text-based, icon-based, and text and icon-based links. Table 8-1 shows an example of three different navigation link format presentations.

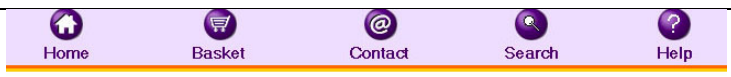
Variation of navigational links	Presentation formats
Home Basket Contact Search Help	Text-based
	Icon-based
	Icon and text-based

Table 8-1 - Example of different presentation format of navigation links

Existing web design guidelines give conflicting advice on the presentation format of the navigation links. Navigation links support the users' navigation through the site, performing two functions: telling the users what information is

held within the site, and helping them to quickly and easily find the information they seek (Public Transport, 2005).

Barde (2002) describes navigation links as aids of how to navigate most effectively the content of a website. Haine (1998) suggested that to be effective and unambiguous a link's label should indicate both a unique reason for selecting the link and the expected results of selection. Table 8-2 shows a summary of web design guidelines for navigation links.

	Authors	Source	Derived from	Suggestions
Text – Based Navigation Links	Detweiler and Omanson (1996)	Ameritech web page user interface standards and design guidelines	User observation	A benefit of using text links is that users with text only and deactivated graphical browsers can see navigation options.
	Spool, Scanlon, Schroeder, Snyder, DeAngelo (1997)	Website usability: a designer's guide	User observation	Users show considerable confusion regarding whether or not images are clickable. This is true even for images that contain words. To avoid any confusion use text based links.
	Koyani and Nalli (1999)	Website design and usability guidelines, national cancer institute	User observation Performance measurement	Text links are more easily recognised as clickable. Text links usually download faster, are preferred by users and should change colours after being affected.
	Nielsen and Tahir (2002)	Homepage design guidelines	Expert inspection	Do not use images when simple text links are clearly differentiable from each other, such in category names. If you find that you need to ponder to come up with an icon for navigation, chances are it is not going to be easily recognisable or intuitions users.
Icon– Based Navigation Links	Van Duyne, Landay, Hong (2003)	The design of sites	Expert inspection	Text navigation links are not quite right for representing transactions. Image navigation links make more sense than text links because they map better with the physical world.
	IBM (2004)	Ease of use: web design guidelines	Expert inspection	Enable users to scan quickly to identify links they want to use. Make sure the images accurately reflect the destination or resulting action of links.
	Horton (1993)	The icon book	User Observation	Icons are useful as reminders Distinguishable icons can aid recognition Icons can save space on screen real estate.

Table 8-2 - Summary of design guidelines regarding navigational links and their presentation format

There are several trade-offs to consider when choosing the presentation format of navigation links. Text-based navigation has been suggested as optimum (Detweiler and Omanson, 1996; Koyani and Nall, 1999; Spool *et al*. 1997; Nielsen and Tahir, 2002) because it enables users to understand the link destinations. In addition, it is accessible for users with text-only browsers. Whereas icon-based navigation has been presented (Van Duyne *et al*. 2003; IBM, 2004, Horton, 1993) as a suitable solution to enable searching and quick identification of the website's links.

As given in the literature the presentation format of navigation links varies and there is no agreement on an efficient way of presenting navigation links on websites. A question that is raised is whether text, icon or both text and icon-based navigation links are more effective for the users when trying to identify a key target link on e-commerce websites such as the shopping cart.

Study 4 aimed to explore how the presentation format of navigation links of e-commerce websites influence the users' visual search performance. A further variable investigated in this experiment was the effect of the location of the target link. The reasoning was that any observations made would not be attributed to the specific location of the target link.

The study addressed the following questions:

- Q1. Does the presentation format of the shopping cart on a web page influence users' visual search?

- Q2. Does the placement of the shopping cart on a web page influence users' visual search?

- Q3. Does the overall design layout of the web pages influence visual search.

8.2 Method

Study 4 consisted of two components: one that investigated user preferences and one that investigated visual search in terms of performance and behaviour. The study goes beyond simple performance metrics that will provide information about the efficiency of the user interface. There is however more to the usability of a user interface than merely the efficiency and effectiveness with which it is used. The subjective preferences and aesthetic impressions of the users also play a very big role in the overall usability (Lindberg and Näsanen, 2003). Therefore Study 4, similalrly to the other studies in this research, has combined several evaluation techniques to explore multiple aspects of the user's behaviour.

The between-participants study using eye tracking collected four types of data:

- *A background questionnaire* to elicit the users' internet experience and typical usage,

- *Pre-session questions* to elicit users' expectations about the use of the shopping cart on e-tail web pages,

- *Eye tracking data* to measure visual search behaviour and performance, and

- *Post-session questions* to address perceptions and preferences about the presentation format and placement of the shopping cart on e-tail web pages.

8.2.1 Study design

The between-participants design used three groups of ten participants each, or 30 participants in total (Table 8-3). Each group was assigned randomly to one of the three conditions:

- *'Text-based'* the shopping cart is presented in written words, text,
- *'Icon- based'* the shopping cart is presented as an icon,
- *'Text and Icon-based'* the shopping cart is presented both in written text and as an icon.

Three e-tail web pages were each modified to conform to each of the three conditions (Table 8-5). To maximise ecological validity within a laboratory experiment (see Chapter 4) existing e-tail web pages were used for Study 4. The presentation order of the stimuli web pages was randomised for each participant.

The study was designed to address one of the issues raised in the literature: that the presentation format of navigation links influences the users' search efficiency. The study compared performance and behaviour on one task across the three exposures and across different presentation format conditions (see Table 8-3):

- 'Please look at the web page and tell me where on the interface you see the shopping cart'.

	Participants	Exposure 1	Exposure 2	Exposure 3	Condition
GROUP A	10	Homepage1a, Homepage2a, Homepage3a appearing in a random sequence for each participant within each group			Text
GROUP B	10	Homepage1b, Homepage2b, Homepage3b appearing in a random sequence for each participant within each group			Icon
GROUP C	10	Homepage1c, Homepage2c, Homepage3c appearing in a random sequence for each participant within each group			Text and Icon

Table 8-3 - Design for Shopping Cart Study

8.2.2 Participants

Study 4 was conducted with 30 volunteers from the staff and postgraduate population of the Open University. Each participant received an honorarium of £10 to acknowledge their time spent in the study. The participants were randomly allocated to one of the three conditions. A detailed presentation of the participants' profiles and their frequency and purpose of internet use can be found in Table 8-4.

Sample Size	Characteristics	Text-based condition	Icon-based condition	Text and Icon-based condition
Total number of participants 30 – 10 in each condition	Age range	23 - 45	23 - 43	22 - 56
	Sex	6 males – 4 females	4 males – 6 females	5 males – 5 females
	Frequency of internet use	3 every day 7 regularly throughout the day	3 every day 7 regularly throughout the day	2 every day 8 regularly throughout the day
	Purpose of internet use	To receive and write e-mail To buy or sell general products online To buy gifts to be delivered to friends or family To buy specialist equipment As an information source As a business medium To chat To listen to music To study online To read newspapers To download music To donate money for charity		

Table 8-4 - Characteristics of participants as elicited from background questionnaires

8.2.3 Stimuli and equipment

E-tail websites were chosen as the domain of interest because:

- Similarly to the other studies in this thesis, it is commonplace, with numerous companies and high profile on the web, and with broad usage across the experience spectrum,
- It is diverse, offering a variety of both services and 'products' for sale in many configurations; and
- It does not have a consistent style.

The selected websites were chosen because:

- they were not leading e-tail websites to decrease the possibility that the participants would have seen them before,
- They shared a similar overall design layout, so that layout was not confounded with presentation format or location of the shopping cart, and
- They could be modified in a way that would enable design elements such as **menu bars** embedded with the target link, to be controlled for.

The selected websites were:

- *Beautiful things* www.beautifulthings.com (e-tail),
- *Prezzy box* www.prezzybox.com (e-tail), and
- *Out of the hat* www.outofthehat.co.uk (e-tail).

Table 8-5 shows screen shots of the stimuli and large images of screen shots can be found in Appendix B.

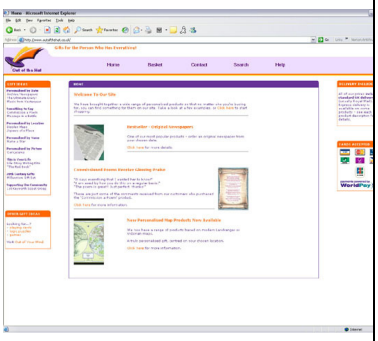
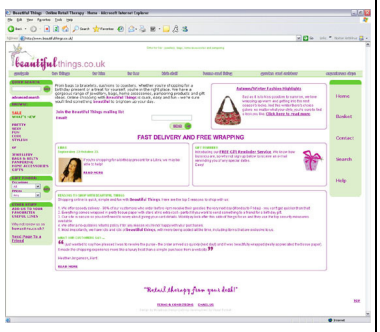
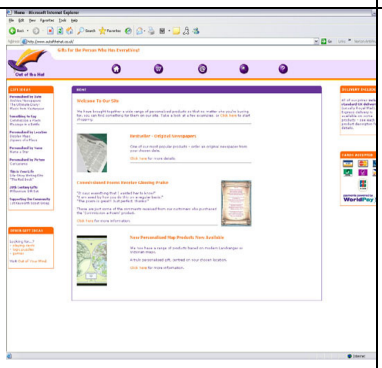
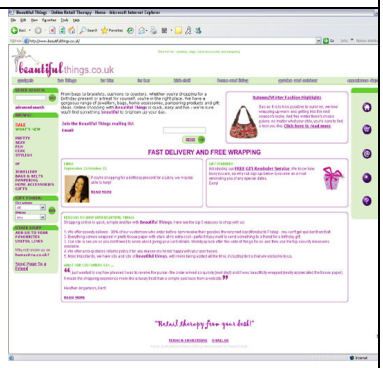
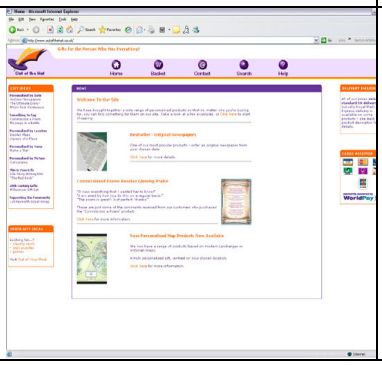
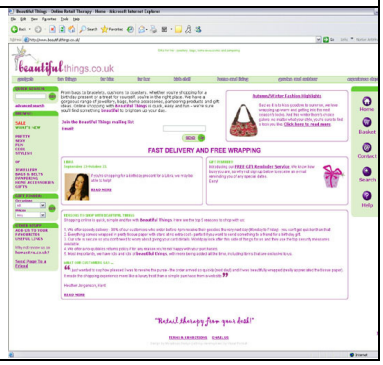
	www.outofthehat.co.uk	www.prezzybox.com	www.beautifulthings.com
Text			
Icon			
Text and Icon			

Table 8-5 Modified homepages of e-tail websites as appearing in three different conditions

Eye movements were recorded using an ASL (Applied Science Technologies) 504 eye tracking remote pan-tilt camera capturing eye movement data at a sample rate of 50 Hz and a minimum fixation threshold of 200 ms. The presentation of the stimuli was controlled by means of the GazeTracker™ software. The screen was viewed by participants from a distance of 55 cm. An infrared light coming from the camera, which was placed under the screen, illuminated the corneal and pupil

reflection. From these two points, the ASL software computed the pupil diameter and line of gaze for each fixation. In other words, this recorded the location of a participant's fixation on the screen. Moreover, it recorded timings of the eye movement data such as time to target fixation (that is the time required to fixate on a given target), along with position and order of the data such as scan path or sequence of fixations. These data were sent via a serial data stream to the GazeTracker™ data analysis software for further manipulation. The GazeTracker™ software allowed the calculation of the time spent looking at predefined areas which are called *Look Zones*, as well as the number of fixations and revisits to the area.

8.2.4 Procedure

The duration of one session including the briefing and **calibration** process was approximately 30 minutes. The session started with a spoken introduction to the participant of the eye tracking equipment and the study. The participant completed a consent form and a background questionnaire which captured age, gender, previous internet experience, and frequency and purpose of internet use. The participant was asked the following scripted questions regarding the presence or absence of images on e-tail websites:

- *Have you done online shopping before?*
- *What kinds of things have you bought?*
- *What is the procedure you go through?*

- *Are there any design elements that you use when shopping online? By design elements I mean features such as a shopping cart.*

- *Where on the interface do you expect to see the shopping cart?*

These questions aimed to collect information about the user's expectations and preferences regarding the presentation format and placement of the shopping cart presence on e-tail websites.

The researcher calibrated the eye tracking camera for the individual. The participants were asked to look at three web pages (Exposures) and complete the same task three times. The presentation order of the web pages was randomised.

The participant was asked to say where on the interface they found the target link in order to indicate that the task was completed, so that the researcher could press the keyboard for the next page to appear. To avoid the researcher's reaction times influencing the data, the eye movement data were used as accurate task completion times.

After the eye tracking session, the participants were asked questions about web pages they had seen during the eye tracking session. The scripted questions were:

- *First of all how did you feel about the tasks?*

- *Was it difficult to complete any of the tasks?*

- *Do you have any preference of the three web pages you just saw?*

- *Any reason for that?*

- *Here I have print outs of more web pages and the ones you just saw. Have a look and tell me what do you think about them?*

- *Which ones do you prefer and why?*

- *Where do you expect to find the shopping cart?*

These questions aimed to collect information about the user's preferences regarding the presentation format and placement of the shopping cart on e-tail websites after the eye tracking session and to compare those expectations and preferences to those elicited by the pre-session questions.

8.3 Data analysis

The analysis of quantitative data focused on two aspects of visual search (Figure 8-1).

1. Statistical analysis of performance, as measured by time to target fixation:

- comparison of all exposures across all conditions

- comparison of each exposure across all conditions

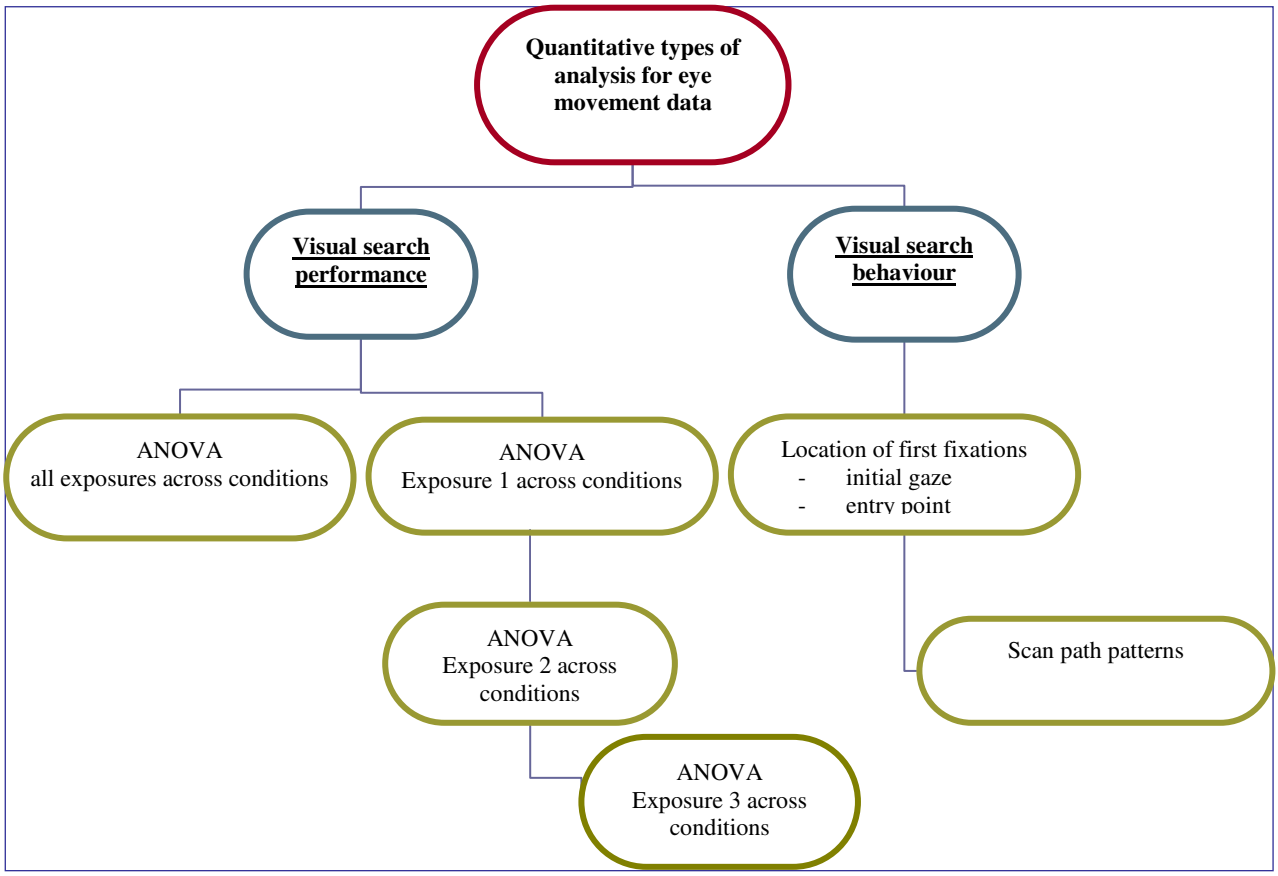


Figure 8-1Diagram of quantitative types of analysis for Study 4

2. mixed quantitative and qualitative analysis of behaviour, as indicated by location of first fixations within *look zones* and location grids, and scan path patterns:

- comparison by exposure
- comparison by condition
- comparison by web page
- comparison by participant

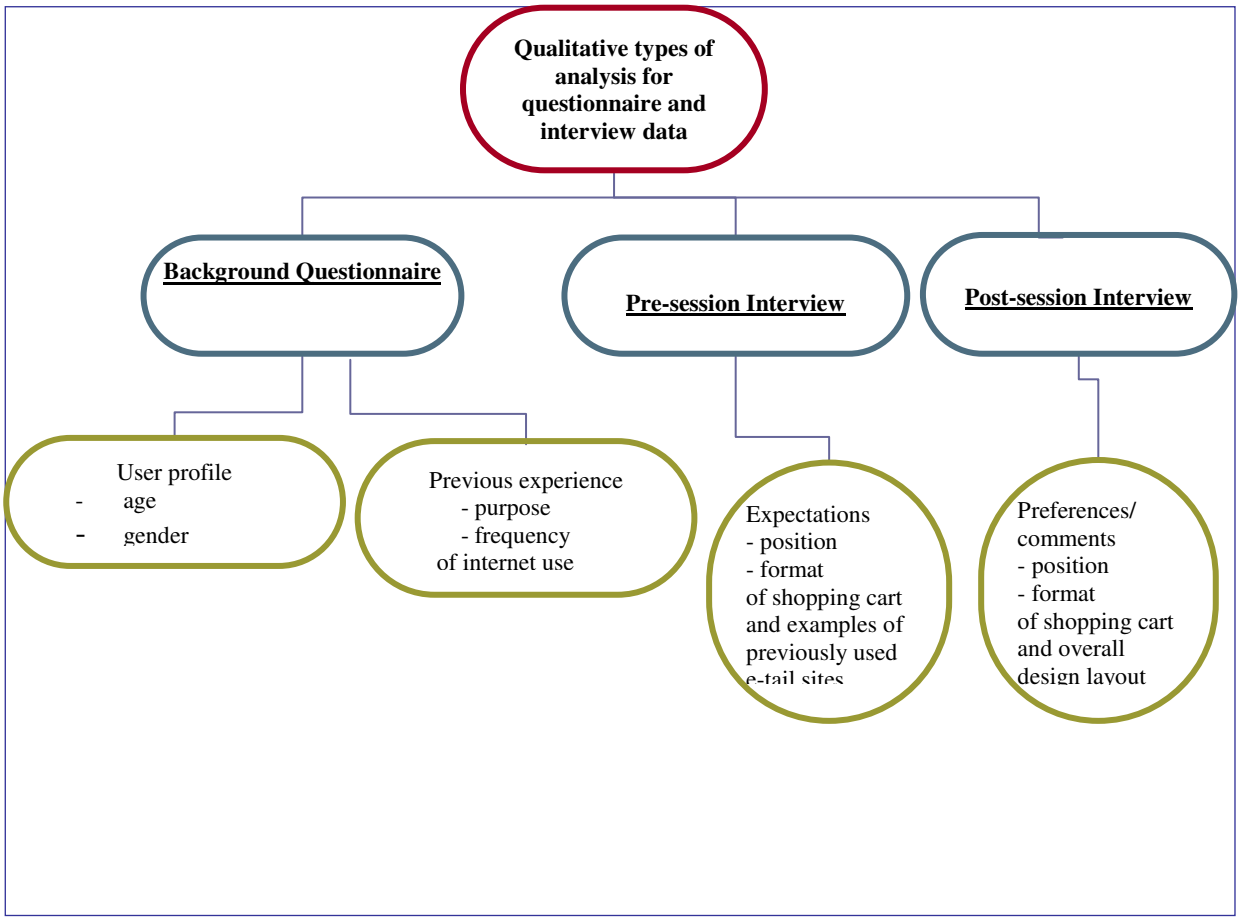


Figure 8-2 Diagram of qualitative types of analysis for Study 4

The analysis of the qualitative data focused on two aspects of the user (Figure 8-2).

1. Analysis of the user profile as gathered by the background questionnaire to elicit information regarding characteristics such as their age and gender along with the purpose and frequency of internet use as indications of their previous experience.

2. Analysis of responses to structured questions during pre and post-session interviews to indicate the users' expectations and preferences regarding the position and presentation format of the shopping cart link. Moreover, the interview sessions aimed to elicit preferences of the design layout of the stimuli web pages and examples of websites previously seen and used for online transactions.

8.4 Results

Visual search performance

The first question Study 4 addressed was whether the presentation format of the shopping cart on a web page influences the users' visual search.

The statistical analysis of the quantitative data enables the exploration of this matter in terms of visual search performance. The metric used for the analysis of the eye movement data is the time to target fixation, capturing the time that the user needed to identify the shopping cart link on the screen.

Comparison of all exposures across conditions

The descriptive statistics of the times to target fixation across the three conditions are shown in Table 8-6.

Conditions	N	Minimum	Maximum	Mean	Std. Deviation
TEXT	10	12.45	38.30	26.58	8.80
ICON	10	9.36	32.47	20.83	6.73
TEXT AND ICON	10	10.31	32.55	17.04	6.84
Total	30				

Table 8-6 Descriptive statistics of sums of time to target fixation (in milliseconds) across aggregated exposures and all conditions

A one-way analysis of variance (**ANOVA**) was calculated on participants' aggregated scores of time to target fixation across the three conditions for all three

exposures. The analysis showed that differences between conditions were statistically significant, F (2,27) = 4,08, p<.05 and a Tukey HSD (Honestly Significant Difference) post-hoc test indicated that there were statistically significant differences between the text-based condition and the icon-based condition and between the text-based condition and the text and icon-based condition.

Comparison of each exposure across conditions

The analysis of all exposures across the three conditions revealed significant differences when aggregating the scores of time to target fixations across the three exposures. In order to examine whether these differences are attributed to any of the three exposures further analysis was conducted to compare each exposure across the three conditions.

Three one-way analysis of variance (**ANOVA**) were calculated on participants' time to target fixation for each exposure across conditions. The analysis showed significant differences only for Exposure 1, F (2,27) = 4.56, p<.05 and a Tukey HSD post-hoc test indicated that there is a statistically significant difference between the text-based condition and the icon-based condition and the text and text and icon-based condition for Exposure 1.

Table 8-7 demonstrates the descriptive statistics for the comparison of times to target fixation across conditions for Exposure 1 and Figure 8-3 show the significant differences across conditions for Exposure 1.

Conditions	N	Minimum	Maximum	Sum	Mean	Std. Deviation
TEXT	10	2.66	19.97	113.64	11.36	5.50
ICON	10	5.41	16.18	85.62	8.56	3.92
TEXT AND ICON	10	1.22	9.61	57.76	5.77	2.35
Total	30					

Table 8-7 Descriptive statistics of sums of time to target fixation (in milliseconds) across three conditions for Exposure 1

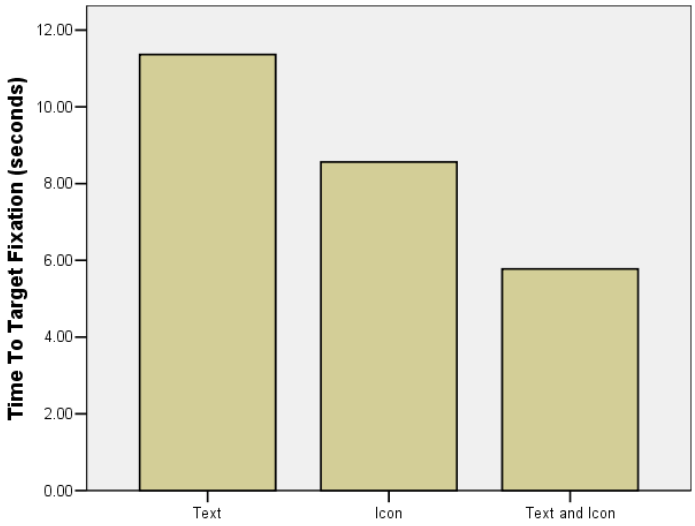


Figure 8-3 Mean of time to target fixation for first task across three conditions for exposure

Discussion of visual search performance results

There were significant differences in performance only for Exposure 1 across the three conditions. It took longer to identify the target link when completing the task in Exposure 1 in the text-based condition when the shopping cart was presented in written text. Given that the task was the same for all three exposures the possibility of the task itself having an impact has been eliminated. The position though of the

shopping cart varied across exposures. For Exposure 1 it was placed on the top navigation bar whereas for Exposure 2 it was placed on the right navigation bar and for Exposure 3 on the left navigation bar. The presentation format, in terms of, size and position within the navigation bar were always the same across exposures. It might be that the position of the shopping cart influenced the significant differences found for Exposure 1. These differences were found only for the text-based condition though that hints that the condition itself might have also an effect in the users' visual search performance. Previous research from Basil (1994) suggests that though people process audio and written words in common ways, they process icons differently. Icons take less mental resources to process; icon processing is more automatic whereas text takes longer to process. The reported results from Study 4 do not fully agree with the results from Basil as significant differences were not found across all exposures for the text-based condition. The results suggest that other factors such as the position of the shopping cart might also be considered as an influential factor. It is expected that the following analysis of the data will contribute further to the discussion.

Visual search behaviour

Three analyses were conducted to explore visual search behaviour:

- Quantitative analysis of the location of initial gazes by location-grid classification
- Quantitative analysis of the location of entry points by design elements classification

- Qualitative analysis of scan path patterns

Location of initial gazes by location grid

Two measures of fixations have been classified:

- '*initial gaze*' where the participant looked less then half a second (50msec) prior to the homepage appearing on the screen in order to examine where users expect to find the shopping cart and

- '*entry point'* as the first fixation within 200msec of the display of the homepage to identify the initial design elements the fixations was drawn to when searching for the shopping cart. The location grid consists of six equal areas on the screen (example in Chapter 4).

Location of entry points by design elements

The GazeTracker™ software allows the classification of *Look Zone*s, areas of interest on the screen, such as **menu bars**, collections of text, or logos. *Look zones* were customised to the stimuli, that is, they were defined individually for each web page. The categorization scheme for the types of *look zones* was a variant of the scheme used by Nielsen and Tahir (2002) to classify design elements on homepages: *Logo, Top* ***menu bar****, Web content, Target menu bar, Right menu bar, Left menu bar, Promotional text,* and *Top menu bar.* An example of the classification of design elements is shown in Figure 8-4.

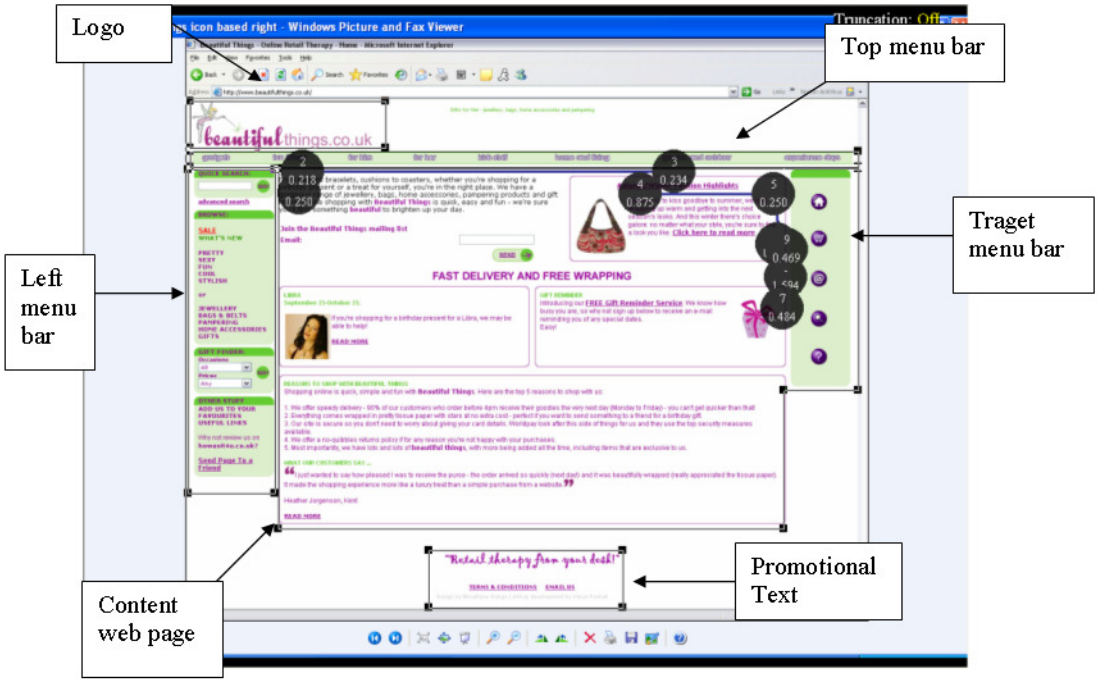


Figure 8-4 Distribution of classification of design elements per look zone on example web page

Scan path analysis of visual search behaviour

A qualitative analysis of the scan paths was conducted in order to classify possible visual search patterns. The classification of different types of scan paths emerged inductively from the scan path data themselves. Four representative scan paths were identified:

Processing was defined quantitatively as ten or more successive fixations on one single design element, such as a **menu bar** or web content (Figure 8-5).

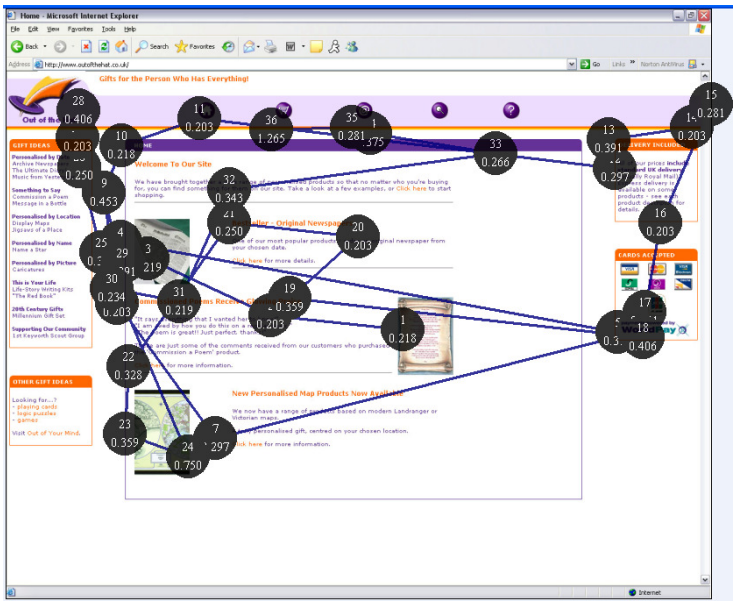


Figure 8-5 Representative scan path, classified as *processing*, superimposed on a stimulus web page in the *icon-based* condition.

Focused scanning was defined quantitatively as very small total numbers of fixations fewer than five fixations for a task. These fixations were often related to the task target (Figure 8-6).

Figure 8-6 Representative scan path classified as *focused scanning* superimposed on a stimulus web page in the *text and icon-based* condition

Random scanning was defined as single fixations on different and non-related design elements. For example, a fixation on text in area A followed on an unrelated image in area C (Figure 8-7).

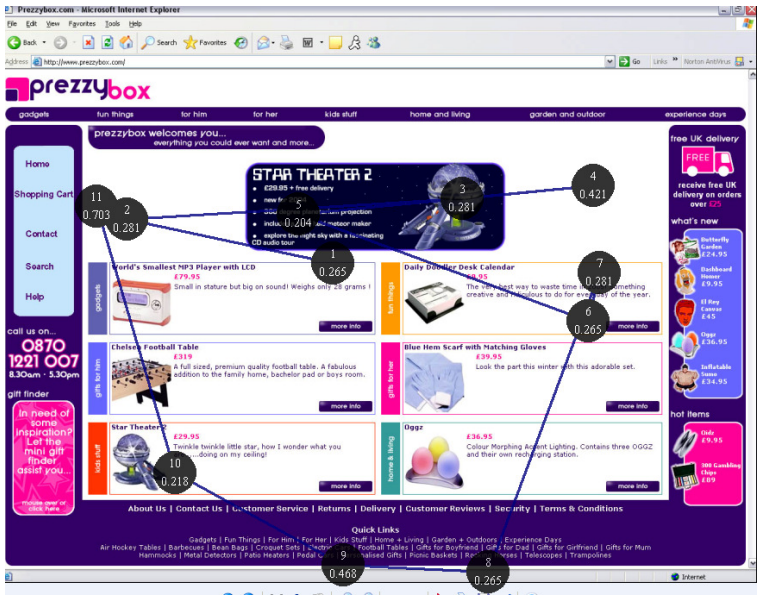


Figure 8-6 Representative scan path classified as *random scanning* superimposed on a stimulus web page in the *text-based* condition

Peripheral scanning: identifying the target link without the scan path 'passing' by it. These fixations were often related to the target task (Figure 8-8).

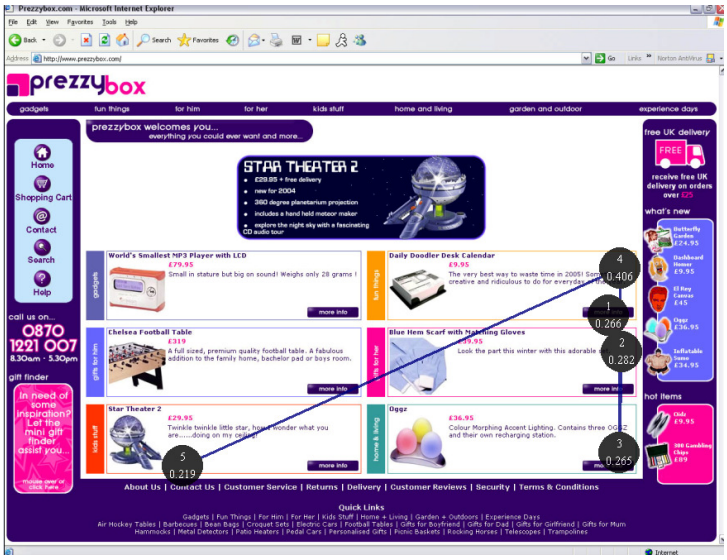


Figure 8-7 Representative scan path classified as *peripheral* superimposed on a stimulus web page in the *text and icon-based* condition

Visual search behaviour

In order to address all three questions of Study 4 a detailed table of the results of each **eye movement-based metric** was created. See in the Appendix B for full length table for all 30 participants across the three exposures and three conditions. The organisation of the data in this matter enabled multiple comparisons of the users' visual search behaviour: a) across conditions, b) across exposures, and c) across participants.

Most frequent and high scores have been reported in a combined format in Table 8-8, with outcomes from all four types of analysis.

Outcomes from different types of eye movement data analysis				
Conditions	**Types of analysis**	**Exposure 1**	**Exposure 2**	**Exposure 3**
Text-based	*Initial gaze*	All initial gazes fell within the top left and top middle of the screen apart from 1	All initial gazes fell within the top left and top middle of the screen (area A and B)	All initial gazes fell within the top left and top middle of the screen (area A and B)
	Entry point	All entry points fell within the web content look zone	The majority of entry points fell within the web content look zone and the rest on promotional text	All entry points fell within the web content look zone apart from one which fell within the top menu bar
	Scan path	Scan path patterns varied across participants: processing, focusing and peripheral scanning	Scan path patterns varied across participants: processing, focusing and peripheral scanning	Scan path patterns varied across participants: processing, focusing and peripheral and random scanning
	Web page	The most shown web page was prezzy box	The most shown web page was out of the hat	The most shown web page was beautiful things
Icon based	*Initial gaze*	All initial gazes fell within the top left and top middle of the screen (area A and B)	All initial gazes fell within the top left and top middle of the screen (area A and B)	All initial gazes fell within the top left and top middle of the screen (area A and B)
	Entry point	All entry points fell within the web content look zone apart from one which fell within the logo look zone	All entry points fell within the web content look zone apart from one which fell within the promotional text look zone	The majority of entry points fell within the web content look zone and the rest on promotional text and logo *look zones*
	Scan path	Scan path patterns varied across participants: processing, focusing and peripheral scanning	Scan path patterns varied across participants: processing, focusing and peripheral scanning	Scan path patterns varied across participants: processing, focusing and peripheral scanning
	Web page	The most shown web page was beautiful things	The most shown web page was prezzy box	The most shown web page was out of the hat
Text and Icon based	*Initial gaze*	All initial gazes fell within the top left and top middle of the screen (area A and B)	All initial gazes fell within the top left and top middle of the screen (area A and B)	All initial gazes fell within the top left and top middle of the screen (area A and B)
	Entry point	All entry points fell within the web content look zone apart from one which fell within the promotional text look zone	All entry points fell within the web content look zone apart from one which fell within the right menu bar look zone	The majority of entry points fell within the web content look zone and the rest on promotional text and top menu bar *look zones*
	Scan path	Scan path patterns varied across participants: processing, random and peripheral scanning	Scan path patterns varied across participants: processing, focusing and peripheral scanning	Scan path patterns varied across participants: focusing and peripheral scanning
	Web page	The most shown web page was prezzy box	The most shown web page was out of the hat	The most shown web page was beautiful things

Table 8-8 Outcomes of different types of analysis across conditions and tasks

Across conditions comparisons

When looking at the outcomes of each type of analysis for the three conditions it becomes evident that:

- Areas that attracted initial gazes were identical across conditions.

For all three conditions 'initial gazes' concentrated on area A and B of the screen. Only one participant's 'initial gaze' was on area C for Exposure 1 of the text-based condition.

There is an overall pattern of *initial gaze*s occurring on the top middle and top left part of the screen despite the different conditions.

The metric of initial gazes aims to capture the user's expectations of where to look for the shopping cart. The examples of websites the participants gave during the pre-session interviews consist of design layouts where the shopping cart is presented on the top right side of the screen. It was therefore anticipated that the location of initial gazes would be found in area C or D of the screen. The unified pattern though of all participants across all conditions suggests otherwise.

- In most cases the most frequent entry point was the web content look zone

Despite varying the presentation format of the target link *entry points* were not found to be influenced. On the contrary, in all conditions the participants seem to follow a common pattern of their scan paths starting off on the largest area covering the screen the *web content*. It is interesting however to mention that entry points

were attracted to distinctive design elements found in the *web content* look zone. For example, the *Prezzy* box web page displays pictures of the products on the stimuli web page as seen by the participants and often entry points were concentrated on the images rather than the text. Likewise, the *beautiful things* web page displays an image of a hand bag on the *web content* look zone, and the *Out of the hat* web page displays images of their products which attracted mostly entry points.

These observations are in line with previous research by Spool (1999) who conducted an eye tracking study on the way people navigate websites. There was no specific task given to the participants to complete. They were asked to freely navigate the sites. The authors found that users' entry points were concentrated in the middle of the screen across several different websites.

The fact that there is a tendency of users to start off their scan path somewhere in the middle of the screen might be due to the fact that in this way there is an equal distance to all other positions on the web page. This is supported by the behaviour of the participants in another study as well as conducted by Stenfors, Moren and Balkenius (2001). They asked users to predict the location of a yellow dot on the screen. It was observed that when the users were not sure at which of the two locations the dot will appear, they often chose to fixate at a position in the middle of the screen between the two expected locations.

- Scan path patterns varied between all three conditions

Table 8-9 demonstrates the distribution of classification of scan path patterns across conditions and exposures.

Scan path visual search behaviour						
Text-based condition	**Exposure 1**		**Exposure 2**		**Exposure 3**	
	Processing	3	Processing	3	Processing	1
	Focused Scanning	3	Focused Scanning	4	Focused Scanning	5
	Random Scanning	0	Random Scanning	0	Random Scanning	1
	Peripheral Scanning	1	Peripheral Scanning	1	Peripheral Scanning	1
Icon-based condition	**Exposure 1**		**Exposure 2**		**Exposure 3**	
	Processing	3	Processing	2	Processing	2
	Focused Scanning	6	Focused Scanning	5	Focused Scanning	5
	Random Scanning	0	Random Scanning	0	Random Scanning	0
	Peripheral Scanning	1	Peripheral Scanning	3	Peripheral Scanning	3
Text and Icon-based condition	**Exposure 1**		**Exposure 2**		**Exposure 3**	
	Processing	0	Processing	1	Processing	1
	Focused Scanning	6	Focused Scanning	4	Focused Scanning	6
	Random Scanning	1	Random Scanning	0	Random Scanning	0
	Peripheral Scanning	1	Peripheral Scanning	3	Peripheral Scanning	1

Table 8-9 Scan path patterns as classified across conditions and exposures

In all conditions the most popular scan path pattern observed was focused scanning apart from the text and icon-based condition where for Exposure 1 and 2 processing was found almost equally popular. This might be influenced by the nature of the task itself which was not exploratory but target orientated. Thus focused processing might have been expected, but what is interesting here is that regardless of the condition and the exposures there is a pattern across participants of similar visual search behaviour.

A new scan path pattern, that has not been observed in our previous studies, has been introduced in Study 4. Peripheral information can elicit extremely rapid shifts of visual attention. At any given moment, peripheral vision accounts for the vast majority of our internal visual representation (McCormick, 1997). Typically when users orient their attention to different locations within the visual field, they execute an overt attention shift, by moving the eyes, lead, or entire body, to align the fovea with the new object of interest. However, while the focus of attention may coincide with the area of the visual field to which the fovea is directed the two are also potentially dissociable 'looking out of the corner of your eye'. Attention can be drawn to a peripheral event even when the observer is not subjectively aware of the source of stimulation. This might help understand how certain users were able to report having seen the target link when in fact their scan path had not passed by the shopping cart.

- Characteristics of the web pages stimuli attracted attention.

Despite the stimuli web pages sharing a common overall design layout, there were some characteristic differences among them, such as the colour combination of text and background, product images, content and so forth. It is worth mentioning though that the navigation bar on which the target link, the shopping cart, appeared was identical in length and size across all exposures and conditions. Moreover, the position of the target link within the navigational bar was identical across both all exposures and conditions. The presentation order of the stimuli web pages was randomised for each participant to avoid any order effects. Because of this there are most frequently displayed web pages for each exposure, although

each participant saw all three stimuli web pages. When looking at the most shown web pages across conditions a relationship was observed in relation to the location of entry points, as discussed in the previous sub section.

Across exposures and across participants comparisons

Despite the task for each exposure being the same, the position of the target link was specific for each exposure. Therefore the shopping cart was placed on the top of the screen for Exposure 1, on the left side of the screen for Exposure 2 and on the right side of the screen for Exposure 3.

When exploring possible patterns of each participant across exposures, only four participants across all three conditions (two in the text-based, one in the icon based and one in the text and icon-based condition) applied an identical visual search behaviour across all three exposures. For example, the location of the initial gaze was area B, the location of the entry point was the *web content* look zone, and the scan path pattern was focused across all three exposures and all three different stimuli web pages. Insights of their profile and examples of their previous experience as captured from the background questionnaire and pre-session interview suggest that these four users were experienced in terms of the frequency and purpose of internet use. They used the internet as part of their job and also had previously completed many transactions online. The consistency of visual search behaviour strategies was not echoed in the other 26 participants.

Pre- and post-session questions

Before and after the eye tracking session the participants were asked a set of open-ended questions regarding:

- previous experience completing an online purchase,
- expectations of the location of the shopping cart on the interface, and
- preferences of the presentation format of the shopping cart.

The post-session interview included the presentation to the participants of print-outs of the web pages in all three conditions. The participants were asked to choose the one they liked the most. Table 8-10 and 8-11 summarises the responses across participants.

Pre-session responses

All of the participants had previously seen e-tail websites, but not all had completed a purchase online. The most frequently visited e-tail websites were E-bay, Amazon, and Tesco. More participants reported having seen the shopping cart on the top right side of the interface. The shopping cart was most preferred when presented as an icon.

Conditions		Pre- session Responses				
		Previous experience using e-tail websites	**Examples of use of websites**	**Procedure of using the shopping cart**	**Expected location of shopping cart**	**Preferences of presentation format of shopping cart**
Text-based	**Group A**	8/10 have made a purchase online 2/10 have not completed a purchase online 10/10 have seen e-tail websites	8/10 E-bay 7/10 Amazon 2/10 Tesco	5/10 look for the shopping cart on the top right corner straight away 3/10 like the shopping cart as they can add items they can take out later 2/10 use the shopping cart just because they have to, as a formality	7/10 Top right 2/10 Bottom right 1/10 Top of page	7/10 icon of trolley or basket 3/10 Icon and text
Icon-based	**Group B**	9/10 have made a purchase online 1/10 have not completed a purchase online 10/10 have seen e-tail websites	10/10 E-bay 4/10 Amazon 2/10 E-buyer 1/10 Tesco	7/10 look for the shopping cart on the top right corner straight away 2/10 like the shopping cart as they can add items they can take out later 1/10 use the shopping cart just because they have to, as a formality	9/10 Top right 1/10 It does not matter	10/10 icon of trolley or basket
Text and Icon-based	**Group C**	8/10 have made a purchase online 2/10 have not completed a purchase online 10/10 have seen e-tail websites	7/10 E-bay 5/10 Amazon 2/10 Tesco 1/10 HMV	7/10 look for the shopping cart on the top right corner straight away 3/10 like the shopping cart as they can add items they can take out later	8/10 Top right 2/10 Bottom right	10/10 icon of trolley or basket

Table 8-10 Consolidated data pre-session questions presenting the frequency of responses to semi-structured questions across participants

Post-session responses

Interestingly, those who said that they preferred seeing the shopping cart as an icon before the eye tracking session changed their preferences after the eye tracking session when shown print-outs of the stimuli web pages. They preferred pages where the shopping cart was presented as both text and icon. One participant said: '*I thought just a trolley or a basket would be fine, but now I can see that the text under it makes it even easier, I like it with both*'. Another participant said: '*Now that I come to think of it, the little symbol might be confusing on its own, I guess personally I am used to it and I know what it means but in general having both the symbol and the text works out well, I must say I prefer this*'.

Prezzy Box and *Out of the Hat* were selected as the least preferred web pages because the design layout was perceived as cluttered.

Conditions		Post- session Responses			
		Difficulty on completing tasks	**Preferences on stimuli web pages**	**Expectations of placement of target link**	**Preferences on conditions**
Text-based	**Group A**	2/10 Confused to find the shopping cart written in text 8/10 No difficulty	5/10 Beautiful things 4/10 Out of the hat 1/10 Prezzy box	8/10 Top right 2/10 top of page	10/10 Text and Icon
Icon-based	**Group B**	1/10 Icons were to small 9/10 No difficulty	6/10 Beautiful things 3/10 Out of the hat 1/10 Prezzy box	9/10 Top right 1/10 Bottom right	10/10 Text and Icon
Text and Icon-based	**Group C**	10/10 No difficulty	5/10 Beautiful things 4/10 Out of the hat 1/10 Prezzy box	8/10 Top right 1/10 Top of page 1/10 Bottom right	10/10 Text and Icon

Table 8-11 Consolidated data post-session questions presenting the frequency of responses to semi-structured questions across participants

The data from the background questionnaire and pre-and post-session questions have enabled a better understanding of the relationship between the users' preferences and visual search behaviour.

8.5 Discussion

The literature on navigation links reveals mixed results but clearly suggests that the presentation of navigation links affects user perceptions.

It is taken into account that the interaction between users and navigation links on the web is a very complex process, made up of many variables such as the size of the links itself, the semantic meaning of the written text or the icons.

Study 4 focussed on two variables only:

- presentation format, and
- position of target link.

A straight-forward search task was chosen to explore the possible effects of these two variables.

Below study questions are revisited to answer them based on the outcomes of the results.

Q1. Does the presentation format of the shopping cart on a web page influence users' visual search looking for the shopping cart?

Significant differences were found only when comparing Exposure 1 across the three conditions. It took significantly longer to find the shopping cart in the text-based condition. When comparing scores of time to target fixation for Exposure 2 and 3 across the three conditions no significant differences were found. Reflecting on previous research, as explained in Chapter 3, it seems that when navigation links are presented in written text it takes longer to identify as text requires more cognitive processes than the identification of icons, but significant differences were found only for Exposure 1. It would be expected then that the third condition where both icon and text appears would require even longer period of times for target identification. It might be the case though the text supports the icon and enables a faster identification of the target link which leads to the possibility of the position of the shopping cart influencing visual search performance in the text-based condition. Further analysis though of the eye movement data and responses to the background questionnaires along with the pre and post-session interviews did not suggest differences across exposures or across conditions. The participants applied similar strategies for all exposures regardless the condition group they participated in.

Q2. Does the placement of the shopping cart on a web page influence users' visual search?

As far as first fixations are concerned there was no evidence that the position of the shopping cart influences visual search. Both measures of location of 'initial gazes' and 'entry points' suggest that different users apply similar initial behaviour in all exposures and all conditions.

As far as preferences are concerned there was a strong preference from all participants of finding the shopping cart on the top right of the screen as that was the position they were used to seeing in on other e-tail web pages they suggested as examples of previous experience. It was only in terms of visual search performance that the position of the shopping cart was found to be a influential factor. In fact for Exposure 1 in the text-based condition scores to target fixation were higher.

Q3. Does the overall design layout of the web pages influence visual search?

The order in which each web page appeared in per exposure was randomised for each participant. This enabled the exploration of possible effects in visual search from a particular design layout.

Despite all three web pages sharing a common general design layout and the **menu bars** where the target link appears on is the same size and each web page has distinctive design characteristics. It seems though that the results do not reveal major differences when comparing visual search behaviour between the web pages. It was only when exploring the location of entry points that distinctive design elements within the *web content* look zone attracted initial visual attention.

But when asking the participants to choose their preferred web page *Beautiful things* collected the highest rates. However, even though participants had claimed to like a certain web page more than the other two, when exploring their visual search no noticeable differences were observed across all three web pages.

Overall, the conclusion to draw from Study 4 is that the user's visual search performance is enhanced by labelling icon-based navigational links. Text-based

links take longer to identify in terms of visual search performance only when the shopping cart was placed on the left **menu bar** of the screen. The presentation format of navigation links, the order of exposures the users performed the task in, and the design layout of the web pages were not found to be influential variables in visual search behaviour. The majority of participants from different condition groups applied a common pattern of initially looking at the centre of the screen carrying out a focused scan path in order to identify the shopping cart.

Future work could explore other variables that might influence visual search performance and behaviour. For example the:

- Colour
- Size
- The wording of the text
- The semantic meaning of the icon of the target link that might play a significant role.

The variation of different tasks could be another way of examining if the results of Study 4 are influenced purely by the specific task chosen.

Chapter 9 Conclusions and further work

My hour at last has come; yet not ingloriously or passively I die, but first will do some valiant deed, of which mankind shall hear in after time.
Homer (900 - 850 BC)

To date, little empirical work has been carried out, or at least reported, in the area of web design guidelines development. This research has taken an empirical, user-centred approach to test existing design guidelines for e-commerce websites. The purpose of this chapter is to review and reflect upon the conclusions drawn from this research, and to propose future research topics and directions.

This chapter starts with a brief overview of the structure of the research. Then the research questions will be revisited, showing how they have been addressed in the research. There follows a brief discussion of directions for further work. The chapter closes with a summary of the conclusions of the reported research.

9.1 Recap of overall research structure

The research studies were discussed in chronological order (Chapters 5-8), because issues that arose in earlier studies influenced the design and analysis of subsequent studies. Despite this inter-connectedness, the studies were independent investigations that explored different subsets of the overarching research questions in different application contexts. Figure 9-1 gives an overview of the studies.

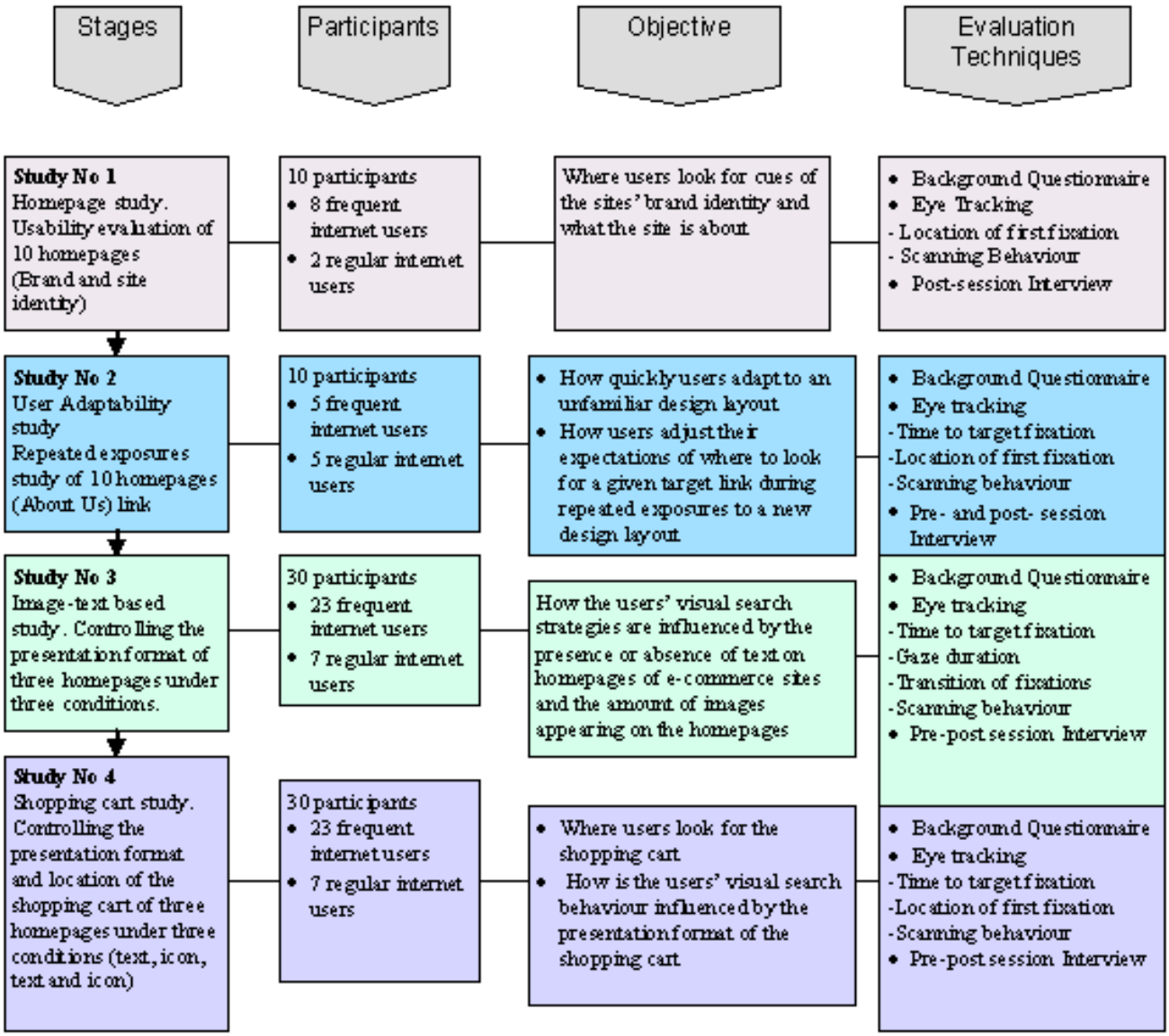


Figure 9-1 Overview of empirical studies

All four reported studies shared a common methodological framework, as explained in Chapter 4, combining eye tracking with more **conventional techniques** such as background questionnaires and pre- and post-session interviews. The analysis was both quantitative and qualitative and took into account performance, visual search behaviour, and user experience, expectation, and

preferences. The intention was to triangulate among sources and analyses in order not only to reveal relationships between design decisions and user performance in the context of e-commerce websites, but also to explicate how different factors shape visual search behaviour.

9.2 Answering the research questions

The overall aim of the research was to test empirically existing web design guidelines for e-commerce websites from a user perspective based on actual user activity. The main research question – What factors influence users' visual search behaviour and performance when they look for specific information on an e-commerce site? – was addressed by focusing on the impact of specific design decisions (such as proportion of images, presentation of key navigational links, location of key information elements) and of expectations and previous experience on users' task performance and visual search behaviour. The main outcomes regarding the influence of both design layout and previous experience for each of the four studies are presented in Table 9-1.

Through a series of focused studies, the research examined a number of factors and articulated what effects on behaviour are associated with those factors that influence visual search behaviour most with respect to finding specific information on an e-commerce website. It related those findings to existing web design guidelines, suggested refinements, and hence showed how eye tracking studies can add value to usability evaluation. Outcomes of the four studies aggregate to contribute insights about two broad issues: where users look for target links, and which design elements attract users' attention most.

Studies	Seeking target information	Which design elements attract attention
Study 1 (Chapter 5)	Previous experience of layout conventions influences initial scanning behaviour. Conventional location of cues for brand identity (such as the logo) on the top left corner or the middle of the page influenced initial fixations.	Layout conventions can establish expectations about the significance of information which lead to misunderstanding. Banners were mistaken as brand identity of the site when located at the top of the screen in distinctive formatting. Those who dislike product images may scan them less thoroughly. Liquid 'design layouts' do not promote completeness of scanning. Scanning tended to be more complete for fixed layouts which fit on a single screen.
Study 2 (Chapter 6)	Users adapt quickly to new layouts if the layouts are self-consistent. However, this adaptation does not completely override the effect of previous experience, which influences where users look for typical information.	Distinctive formatting is effective only if used sparingly and strategically. Distinctive formatting of design elements such as emboldening and position of target link within navigational bar improves performance – but only if the formatting is truly distinctive.
Study 3 (Chapter 7)	Users' deployment of visual search strategies is influenced by the nature of the task.	The proportion of images on the screen does not itself influence performance, nor does the absence of images. However, the layout does influence visual search behaviour, with a good balance of images and text leading to more orderly scanning. Dominant features such 'text' or 'images' create a 'pull-down' effect. If those features are pertinent to the task, then performance is enhanced. However, if they are not, then the user may be distracted.
Study 4 (Chapter 8)	Users' sensitivity to layout conventions is influenced by the nature of the task (for example recognising versus reading and comprehending). The placement of navigational links was not an influential factor in users' visual search behaviour, regardless of the users' previous experience (such as being used to finding the shopping cart on the top right of the screen). Peripheral vision may be exploited during recognition tasks.	Key navigational links are best presented in redundant formats. Users found the shopping cart most quickly when it was presented as both icon and text. It took longer to find the shopping cart when presented as text only.

Table 9-1 Outcomes of studies relating to the effect of previous experience and design layout

9.2.1 Seeking target information

When exploring the influence of previous experience on task performance, specifically where users expect to find target information on the user interface, the four studies suggested the following:

- *Expectation and first impressions:* Users look on the top left and top middle of the homepage for clues of the brand identity of the website (Study1). This is influenced by the users' previous experience of interacting with websites on which the logo and name of the company appear on the top left and top middle of the page. Therefore it is concluded that placing design elements regarding the brand and main purpose of the site in the expected locations on the user interface makes users more likely to recognize brand identity more effectively.

- *User adaptability:* Users 'learn' where to look over repeated exposures, but the effect of their previous experience overrules their quick adaptation to a specific placement of the design link. What they have seen on other websites was found to be a stronger effect (Study 2). This might release designers from design convention – but only if they can keep the user on the site through several pages. A strong consistent design is more likely to enhance the user's performance rather than just meeting the users' expectations.

- *Importance of task context:* The purpose, relevance and familiarity of the task itself influence users' behaviour. For example, when completing a very familiar task, users perform faster (Study 3). In another example,

the placement of navigational links was not an influential factor in users' initial scanning, regardless of the users' previous experience (Study 4). This is in contrast to the finding regarding scanning for brand identity, for which performance was influenced by placement (Study 1). Hence, the users' sensitivity to design layout depends on the task context – on which sort of information they are seeking and whether they expect to read and comprehend, or just to recognise – and design layout is more important for some tasks than others.

9.2.2 Which design elements attract attention

When exploring the effects of the design layout and in specific which user-interface design elements attract users' attention the most, the four studies suggested the following:

- Layout conventions can establish expectations about the significance of information which lead to misunderstanding. Banners on homepages might be mistaken as the brand identity of the website depending on their size and placement on the interface (Study1).

- The presentation of product images on homepages can result in inefficient scanning for users who dislike images (Study 1).

- 'Liquid design layouts' do not enhance completeness of scanning (Study 1). Scanning tended to be more complete for fixed layouts which fit on a single screen.

- Distinctive formatting of design elements such as emboldening and position of target link within the navigational bar improves performance, but only if the formatting is sufficiently distinctive (that is, it is the distinctiveness that matters, not emboldening *per se*) (Study 2).

- The proportion of images on the screen does not itself influence performance, nor does the absence of images (Study 3). However, the layout does influence visual search behaviour, with a good balance of images and text (in this study, five to 15 per cent images) leading to more orderly scanning.

- Dominant features such 'text' or 'images' create a 'pull-down' effect, meaning that they draw attention (Study 3). If those features are pertinent to the task, then performance is enhanced. However, if they are not, then the user may be distracted. Dominant features can influence visual search behaviour, resulting in more complex scanning patterns when they are not relevant to the task.

- When a navigational link, such as the shopping cart, is presented textually it takes longer to identify. Such target links are identified more quickly when presented as both an icon and in text (Study 4). Peripheral vision may be exploited during recognition tasks such as finding a familiar navigational link.

9.3 Further directions

Several issues arose in this doctoral research that warrant further investigation. The combination of evaluation techniques used as indicators of usability has successfully contributed to the analysis of existing web design guidelines for e-commerce sites. However, the richness of the data collected by the different techniques suggests further possible ways of analysing the data, depending on the various aspects of interest. Suggestions of further directions are listed below:

Realistic tasks

According to the results, visual search behaviour was task sensitive. The tasks given to the participants were of two kinds:

- 'exploratory' (that is for example to have a look at the web page and tell me what this site is about), and
- 'target oriented' (that is for example to find the shopping cart and tell me where on the screen you can see it).

When users search for a tool, menu item, icon and so forth on the interface, they often do not have a good representation of the target. Most of the literature in visual search starts with the participant knowing the specific target. There is a need to conduct more basic research in visual search when the target is not completely known. A more realistic search task is to look for a tool that will help you do a specific task, without having seen the tool before. For example, when capturing the experience of a user when making an online transaction on a website they have not used before.

Advanced eye tracking devices

The eye tracking devices used in the research studies restricted the participants from making major head movements, and users who wear spectacles for viewing computer displays were excluded from participating. It is promising that developers of eye tracking devices are making great progress in reducing such physical constraints (for example Tobii Ltd). Enabling the collection of eye movement data in natural settings, such as the work place, and without constraint on the participant's movement, will add to the external and ecological validity of eye tracking studies in usability evaluation.

Different user groups

This research involved only internet-literate users. It is important to consider representative samples of different user groups, such as users who are not familiar with the internet. As was discussed earlier, previous internet experience was found to be an influential factor in the users' expectations of where to find specific design elements. Exploring how users with less internet experience search visually on e-commerce websites is a matter for future research.

Multiple pages –complex tasks

The stimuli used in the reported studies were static web pages. The participants were asked to look at single web pages, each from a different website, and were therefore not allowed further interaction within any given website. Future research should explore the effect of more complex tasks across multiple web pages of the same website.

Relationship between task and scanning behaviour

The studies highlighted the likely importance of the nature of the task on visual search behaviour, with different results regarding the impact of layout conventions for brand identification (Study 1) from those for finding the shopping cart (Study 4) suggesting that 'recognition' strategies may be different than 'reading' strategies. Further, the evidence that peripheral vision is exploited in some tasks (shopping cart recognition, Study 4) raises questions about the relationship between tasks and visual search behaviour – clearly a matter for further, detailed study.

The observed use of peripheral vision poses challenges for the application of eye tracking, because users were able to perform accurately with regard to things on which they did not fixate, and hence the relationship between visual behaviour and cognition was indirect. Hence, more caution is required in drawing inferences from eye tracking data alone. However, the observation illustrates the power of the combined techniques to reveal such surprises and account for them – and hence demonstrating a new role for eye tracking in the context of this combined approach.

Relationship between performance and scanning behaviour

While scanning behaviour and visual search performance have been analysed as two separate aspects in this research, further exploration of the relationship between scanning behaviour and performance might answer questions such as: Are some visual search strategies more efficient, and, if so, under what conditions? Can they be taught? For example, airline pilots are taught systematic scanning behaviours in order to ensure that they take in changes in the airplane's state. Accidents have

resulted from scanning failure, for example when a pilot failed to notice that the automatic pilot had been engaged accidentally (Wickens *et al.* 2004).

Moreover, there is a need to explore what it means when there are differences in visual search behaviour but no significant difference in performance: is the task simply insufficient to reveal performance differences? Or are there other important differences (such as satisfaction, recall, preference) that should be taken into account? Given the e-commerce context, and the recognition that customer preference, satisfaction and retention are perhaps as important as task performance, future research should focus on this matter.

New analysis methods

The qualitative analysis of scan path patterns in Study 3 and 4 provided insights in the interpretation of visual search behaviour. However, such data analysis is labour-intensive. This could be ameliorated by software which could help identify and classify different visual search strategies at different levels of granularity, and which could aggregate scanning patterns across participants and also across different stimuli. Being able to collate such data automatically with the use of software would make it possible to conduct larger studies involving longer observation periods and more complex tasks.

Relating eye tracking to other domains

While eye tracking has been applied in the domain of e-commerce in this research, it could also be applied in other media and marketing research. This research demonstrates ways of complementing eye movement data with data from other techniques. Such a methodological approach could be used for example, in

marketing research for the exploration of the customer in real settings rather than in online stores only.

9.4 Contributions of the thesis

Previous research on website evaluation provides a few examples where eye tracking has been used as an indicator of usability (examples in Chapter 3), but most studies have focused on the study of the effectiveness of eye movement-based metrics. When dealing with eye movement data it is crucial to know what metrics to use and what information can be reached by applying them. In this sense the focus on the use of eye tracking for websites evaluation has been extremely valuable for forthcoming research. Knowing what metrics to use to identify cause and effect has advanced and proven the potential of eye tracking in the field of usability testing. This research has employed knowledge about the appropriate **eye movement-based metrics** to use from past work but also suggests a more advanced methodological approach where eye movement data are supported with data from other techniques (Chapter 4). Only a few other studies have demonstrated the benefits of combining eye tracking with other techniques such as, for example, log event recording (Narayanan and Schrimpsher 2000).

In this research eye tracking revealed the visual search behaviour. The interviews connected the visual search behaviour to user perception and experience. As demonstrated in the conclusions of the four studies of this research eye tracking can be used to mediate contradictions between guidelines. For example, it can expose trade-offs between competing factors, such as expectation and salience (Study 1 and Study 4), layout and task (Study 3), long-term and short term

experience (Study 2). Moreover, it was observed that each technique, used in this reported research (see Chapter 4), was able to detect different aspects of the user behaviour, which would not have been possible by only applying one single technique. For example, when asking the participants to report where they thought they looked on the screen while completing a task, in many cases their responses were different from what they actually looked at as shown by the eye movement data. As eye movements are so rapid and happen subconsciously it is difficult to report where one has looked at. Therefore, eye tracking can capture this fine detail in an objective way.

Overall, the application of the combined methodology revealed aspects of the design of web pages that had not been addressed by existing web design guidelines. The detailed information of the shifts of visual search behaviour and performance on the stimuli web pages gave an insight of which design elements attract attention and where on the user-interface users expect to find specific information about the brand of the company and what they can do within the site. As discussed earlier the majority of existing web design guidelines has been derived has been derived by personal reviews and anecdotes without reporting the involvement of human participants. This research tested and elaborated the existing web design guidelines by capturing visual search behaviour and performance. The findings are expected to benefit HCI designers, as they will learn about the potential use of eye tracking for usability evaluations of products, and also web designers, who may be able to apply the user-centred web design guidelines and also structure the user interface effectively to influence user's visual search behaviour.

References

Antes, J. R. (1974). The time course of picture viewing. *Journal of Experimental Psychology*, *103*, pp. 62-70.

ASL (Applied Science Laboratories) Model 504. http://www.a-s-l.com/model504.htm last visited 29th April, 2005.

Barde, A. N. (2002). *Shaping web usability interaction design in context*. Addison-Wesley, Boston, MA.

Barrick, J., Maust, B., Spyridakis, J.H., Eliot, M., Wei, C., Evans, M., and K. Mobrand. (2004). A Tool for Supporting Web-Based Empirical Research: Providing a Basis for Web Design Guidelines. *In Proceedings of IEEE International Professional Communication Conference*, pp. 189-193.

Barnard, J. M. and **Barnard, P.** (1995), The case for supportive evaluation during design, *Interacting with Computers*, *7*, (2), pp.115-143.

Basil. M. D. (1994). Multiple resource theory. Application to television viewing. *Communication Research*, 21 (*2*), pp. 177-207.

Beier, B., and Vaughan, M. W. (2003). The bull's-eye: A framework for web application user interface design guidelines. *In Proceedings of CHI 2003*, New York, pp. 489-496.

Benel, D., Ottens, D., and Horst, R. (1991). Use of an eye tracking systems in the usability laboratory. *In Proceedings of the Human Factors Society 35th Annual Meeting*, California, USA, pp. 461-465.

Bernard, M., Larsen, L. (2001). What is the best layout for multiple-column web pages? Usability News bulletin 3.2. Wichita University. http://psychology.wichita.edu/surl/usabilitynews/3S/layout.htm, last visited, 21st October, 2005.

Beyer, H., and Holtzblatt, K. (1998). *Contextual design: defining customer-centered systems*. Morgan Kaufmann, San Francisco.

Bhatti, N., Bouch, A., and Kuchinsky, A. (2000). Integrating user-perceived quality into web server design, *Computer Networks*, *33*, pp.1-16.

Booher, H. R. (1975). Relative comprehensibility of pictorial information and printed words in proceduralised instructions. *Human Factors*, *17*, pp. 266-277.

Borges, J., A., Morales, I., and Rodriguez, N. J. (1996). Guidelines for Designing Usable World Wide Web Pages, *In Proceedings of Companion of CHI 96 Computer-Human Interaction Conference*, Vancouver, Canada. pp. 277-278.

BPS (The British Psychological Society). (2002). Code of Conduct, Ethical Principles and Guidelines. www.bps.org.uk/documents/Code.pdf. Last visited 20th October, 2005.

Brinck, T., Gergle, D and Wood, Scott D. (2002). *Usability for the Web: Designing Web Sites that Work*. San Francisco: Morgan Kaufmann.

Byrne, M., D. (1993). Using icons to find documents: Simply is critical. *In Proceedings of INTERCHI*, Amsterdam, The Netherlands, pp.446-453.

Byrne, M. D., Anderson, J. R., Douglas, S. and Matessa, M. (1999). Eye tracking the visual search of click-down menus. In Proceedings of CHI 99, Pittsburgh: ACM Press, pp.15-20.

Campbell, D. and Stanley, J. (1963). Experimental and Quasi- experimental designs for research on teaching. In Gage, N. (Eds.), *Handbook of Research on teaching*. Rand McNally Company, Chicago, pp. 171-216.

Cao, M. (2002). Evaluating e-commerce web site design: a customer perspective. Design Sciences Institute, *In Annual Meeting Proceedings*, pp. 1186-1191 http://www.sbaer.uca.edu/Research/2002/dsi/papers/224.pdf . Last visited 12th October, 2004.

Card, K., Mackinlay, J. D. and Shneiderman, B. (1999). *Readings in Information Visualization: Using Vision to Think*, Morgan Kaufmann, California , pp. 1-34.

Carroll, J. (1998). Models and Theories of Human. Computer Interaction, CS5724, Virginia Polytechnic and State. University, Blacksburg, VA, USA.

Chatman, B. (2002). Exposing customer experience flaws, Technical report, Forrester Research.

Chisholm, W., Vanderheiden, G. and Jacobs, I. (1999). Web Content Accessibility Guidelines 1.0, http://www.w3.org/TR/WCAG10/WAI-WEBCONTENT-19990505/. Last visited, 8th June, 2004.

Clark, **H.** (1916). Visual imagery and attention: an analytical study. *American Journal of Psychology, 27*, (4), pp. 461-492.

Cohen, **N.** (1999). Greening the Internet: ten ways e-commerce could affect the environment and what we can do. Information Impact Magazine, October http://www.cisp.org/imp/october_99/10_99cohen.htm . Last visited, 8th June, 2005.

Cowen, L., Ball, L.J., and Delin, J. (2002). An eye movement analysis of web page usability. *In Proceedings of the BCS-HCI 2002*. People and Computers XVI-Memorable yet invisible, London, UK, pp. 317-335.

Crosby, M., and Peterson, W., W. (1991). Using eye movements to classify search strategies. *In Proceedings of the Human Factors Society 35th Annual Meeting*, *Vol 2*, San Francisco, CA, pp. 1476 – 1480.

Crowe, E.C. and Narayanan, N.H. (2000). Comparing interfaces based on what users watch and do. *In Proceedings of (ETRA) Eye Tracking Research and Applications Symposium*, Florida, USA, pp.29-36.

DeSanctis, G. (1984). Computer graphics as decision aids: Directions for research. *Decision Sciences*, *15*, pp.463-487.

Detweiler, M., C. and Omanson, R., C. (1996). Ameritech Web Page User Interface Standards and Design Guidelines www.ameritech.com last visited 1st June, 2004.

Dix, A., Finlay, J., Abowd, G. D., and Beale, R. (2004). *Human –Computer Interaction*. Third edition, Prentice Hall, England, UK.

Dix, A., Finlay, J., Abowd, G. D., and Beale, R. (1998). *Human-Computer Interaction*, Second edition, Prentice Hall, England, UK.

Dodge, R., and Cline, T. S. (1901). The angle velocity of eye movements. *Psychological Review, 8*, 2, pp. 145-157.

Drury, J., J. Scholtz and Yanco, H.A. (2003). Awareness in Human-Robot Interactions. In Proceedings of the IEEE Conference on Systems, Man and Cybernetics.

Dutt, A., Johnson, H., and Johnson, P. (1994) Evaluating Evaluation Methods. *In Proceedings of the conference on People and computers IX*, Glasgow, Cambridge University Press, pp. 109-121.

Egger, F. N. (2001). Affective Design of e-commerce user-interfaces: How to maximise perceived trustworthiness. *In Proceedings of CAHD: Conference on affective human factors design*. New York, NY: ACM Press, pp.317-324.

Ehret, B. D. (2000). Learning where to look: The acquisition of location knowledge in display-based interaction. Doctoral Dissertation, George Mason University, U.S.A.

Ehret, B. D. (2002). Learning where to look: location learning in graphical user interfaces. *In Proceedings of CHI*, Minnesota, USA , pp. 211-218.

Ellis, S., Candrea, R., Misner, J., Craig, C. S., Lankford, C. P. and Hutchinson, T.E. (1998). Windows to the soul? What eye movements tell us about software usability. *In Proceedings of the Usability Professionals' Association Conference*, Arizona, USA, pp. 151-178.

EPS IT Inc. (2005). Accessed http://www.epsltd.com/clients/statistics.asp, last visited, 3rd October, 2005.

Fang, X., and Salvendy, G. (2003). Customer-centered rules for design of e-commerce Web sites. *In Communications of ACM 46* (12), pp.332-336.

Faraday, P., and Sutcliffe, A.G. (1996). An Empirical Study of Attending and Comprehending Multimedia Presentations. *In ACM Multimedia*, pp. 265-275.

Fitts P. M. (1954). The Information capacity of the human motor system in controlling the amplitude of movement. *Journal of Experimental Psychology, 47*, pp. 381-391.

Fitts, P. M., Jones, R. E., and Milton, J. (1950). Eye movements of aircraft pilots during instrument landing approaches. *Aeronautical Engineering Review 9* (2), pp. 24-29.

Fleming, J. (1998). WEB navigation: designing the user experience, O'Reilly, CA, USA.

Flemisch, F. O., and Onken, R. (2000). Detecting usability problems with eye tracking in airborne battle management support. NATO RTO HFM. *Symposium on Usability of Information in Battle Management Operations*, Oslo, Norway.

Fu, L., Salvendy, G., and Turley, L. (1998). Who finds what in usability evaluation? *In Proceedings of the Human Factors and Ergonomics Society 42nd Annual Meeting*, Santa Monica, CA: HFES, pp. 1341-1345.

Gilliland, A. R. (1921). Photographic methods for studying reading. *Visual Education, 2*, pp. 21-26.

Goldberg, J. H. and Kotval, P.X. (1999). Computer interface evaluation using eye movements: methods and constructs. *International Journal of Industrial Ergonomics 24*, pp. 631-645.

Goldberg, J. H. and Kotval, P.X. (1998). Eye movements and interface component grouping: an evaluation method. *In Proceedings of the human and ergonomics society 42nd annual meeting,* Santa Monica, CA, pp. 486-490.

Goldberg, J. H. and Winchansky, A. M. (2003). Eye tracking in usability evaluation. *In The Mind's Eyes: Cognitive and Applied Aspects of Eye Movements*. R. R. J. Hyona, and H. Deubel, Elsevier, Oxford pp. 493-516.

Goldberg, J. H., Stimson, M. J, Lewenstein, M., Scott, N. and Wichansky, A. M. (2002). Blink response, visual attention, and the www: Eye tracking in web search tasks: design implications. *In Proceedings of ETRA*, Louisiana, USA, pp. 51-58.

Goldthwait, C. (1933). Relation of eye movements to visual imagery. *American Journal of Psychology.*

Granka, L., Hembrooke, H. Gay., and Feusner, M. (2004). *Correlates of Visual Salience and Disconnect*. Unpublished research report, Cornell University Human-Computer Interaction Lab http://www.stanford.edu/~granka/research.html last visited 29th April, 2005.

Granka, L., Joachims, T., and Gay, G. (2004). Eye-Tracking Analysis of User Behaviour in WWW Search. *In Proceedings of SIGIR,* Shefield, UK, pp. 478-479.

Haine, P. (1998). Five most serious web design errors. www.hp.com last visited 1st June, 2004.

Hammersley, M. (1999). *Researching school experience: ethnographic studies of teaching and learning.* Palmer Press, London.

Hansen, J.P. (1990). The use of eye mark recordings to support verbal retrospection in software testing. *Acta Psychologica, 76*, pp. 31-49.

Hansen, D. W. (2003). Comitting eye tracking. Ph.D. thesis, IT University of Copenhagen.

Harris, R. L., and Christhilf, D. M. (1980). What do pilots see in displays? *In Proceedings of the 24th Annual Meeting of the Human Factors Society*. G. Corrick, E. Hazeltine, and R. Durst (Eds.) Santa Monica, CA.

Hartridge, H., Thompson, L.C. (1948) Methods of investigating eye movements, *British Journal of Ophthalmology, 32*, pp 581-591.

Hartson, H.R., Andre, T.S., Williges, R.C. (2000). Criteria For Evaluating Usability Evaluation Methods. *International Journal of Human Computer Interaction 15*, (1), pp. 145-181.

Hartson, H. R., Castillo, J. C., Kelso, J., Kamler, J.,and Neale, W. C. (1996). Remote Evaluation: The Network as an Extension of the Usability Laboratory. *In Proceedings of CHI Conference on Human Factors in Computing Systems* ACM, New York, pp. 228-235.

Hendrickson, J.J. (1989). Performance, preference, and visual scan patterns on a menu-based system: implications for interface design. *In Proceedings of the ACM CHI'89 Human Factors in Computing Systems Conference*. ACM Press, CA USA, pp. 217-222.

Herschiag, M. (1998). Shopping the Net. Time Magazine.

Hornof, A. J. and Halverson, T. (2003). Cognitive strategies and eye movements for searching hierarchical computer displays. *In Proceedings of CHI 2003*, ACM Press, pp. 249-256.

Horowitz, T.S., and Wolfe, J.M. (1998). Visual Search Has No Memory. *Nature, 357*, pp. 575-577.

Horton, W. (1993). *The icon book: Visual symbols for computer systems and documentation*. John Wiley, New York.

Hurst, M. & Terry, P. (2000). *The dotcom survival guide*. Unpublished report, Creative Good, New York, NY.

IBM (WebDesignGuidelines) (1999).

http://www3.ibm.com/ibm/easy/eou_ext.nsf/Publish/572 .Last visited 26th Feb. 2004.

Instone, K. (1997). Usability Heuristics for the Web http://www.webreview.com/1997/10_10/strategists/10_10_97_2.shtml . Last visited on 2nd July, 2003.

Ivory, M.Y., Sinha, R. R. and Hearst, M.A. (2001). Empirically validated web page design metrics. *In Proceedings of the SIGCHI*, Washington, USA, pp. 53-60.

Jacob, R., and Karn, K. S. (2003). Eye Tracking in Human-Computer Interaction and Usability Research. In *The Mind's Eyes: Cognitive and Applied Aspects of Eye Movements*. R. R. J. Hyona and H. Deube (Eds.). Elsevier, Oxford pp. 572-605.

Jarvenpaa, S. L., and Todd, P.A. (1997). Consumer reactions to electronic shopping on the World Wide Web. *International Journal of Electronic Commerce*, M.E. Sharpe, INC, No 2.

Jones, W. P. and Dumais, S. T. (1986). The spatial metaphor for user interfaces: experimental tests of reference by location versus name. *ACM transactions on Office Information systems, 4*, (1), pp. 42-63.

Josephson, S. and Holmes, M.E. (2002). Visual attention to repeated internet images: testing the scan path theory on the World Wide Web. *In Proceedings of ETRA*, Louisiana, USA, pp. 43-51.

JRuler Pro. Pixel measure Software http://www.spadixbd.com/jruler/ .Last accessed 2nd August, 2004.

Just, M. A. and Carpenter P. A. (1976). Eye Fixations and Cognitive Processes. *Cognitive Psychology 8*,(9), pp. 441-480.

Just, M.A. and Carpenter, P.A. (1975). The role of eye fixation research in cognitive psychology. *Behaviour research methods & instrumentation. 8*, pp. 139-143.

Just, M. A. and Carpenter, P. A. (1980). A theory of reading: From eye fixations to comprehension. *Psychological Review, 87*, pp. 329-354.

Kotval, X.P., and Goldberg, J.H. (1998). Eye Movements and Interface Components Grouping: An Evaluation Method. *In Proceedings of the 42nd Annual Meeting of the Human Factors and Ergonomics Society*, Santa Monica: HFES, pp. 486-490.

Koyani, S. and Nall, J. (1999). Web Site Design and Usability Guidelines, National Cancer Institute, Bethesda, MD. http://usability.gov/guidelines/navigation.html#two. Last visited 1st June, 2004.

Krug, S. (2000). *Don't make me think*. Indianapolis: New Riders.

Kujala, S. (2003). User Involvement: a review of the benefits and challenges. *Behaviour and Information Technology, 22*, 1, pp. 1-16.

Law, L., and Hvannberg, E.T. (2002). Complementarity and convergence of heuristic evaluation and usability test: a case study of universal brokerage platform. *In Proceedings of second Nordi Conference on Human Computer Interaction*, Denmark, pp. 19-23.

Lewenstein, M., Edwards, G., Tatr, D., and De Vigal, A. (2000). The Stanford Poynter Project. Available at http://www.poynter.org/eyetrack2000/ last visited 29th April, 2005.

Limbourg, Q., and Vanderdonckt, J. (2002). Comparing Task Models for User Interface Design, Chapter 6, in Diaper, D. and Stanton, N. (Eds.), in *The Handbook of Task Analysis for Human-Computer Interaction*, Lawrence Erlbaum Associates, Mahwah.

Lindberg, T., and Näsanen, R. (2003). The effect of icon spacing and size on the speed of icon processing in the human visual system. *Displays 24*, pp. 111-120.

Loftus, G. R. and Mackworth, N.H. (1978). Cognitive determinants of fixation location during picture viewing. *Journal of Experimental Psychology: Human perception and performance, 4*, 4, pp. 565-572.

Mackay, W.E. and Fayard, A.L. (1997) HCI, natural science and design: A framework for triangulation across disciplines. *In Proceedings of Designing Interactive Systems, DIS'97*, ACM: Amsterdam, the Netherlands, pp. 223-234.

Mackworth, N.H. and Morandi, A. J. (1967). The gaze selects informative details within pictures. *Perception and Psycholinguistics, 7*, pp.173-178.

Mackworth, N.H. and Thomas, E.L. (1962). Head-mounted eye-marker camera. *Journal of the Optical Society of America. 52*, pp. 713-716.

Marion, S. D., Touchette, P. E. and Sandman, C. A. (2003). Sequential analysis reveals a unique structure for self-injurious behaviour. *American Journal on Mental Retardation: 108*, 5, pp. 301–313.

McCarthy, J. Sasse, M.A. and Riegelsberger, J. (2003). Could I have the menu please? An eye tracking study of design conventions. *In Proceedings of BCS-HCI*, Bath, UK, pp. 401- 414.

McCormick , P. (1997). Orienting attention without awareness. *Journal of Experimental Psychology: Human perception and performance, 23*, pp. 168-180.

Minocha, S., Dawson, L., Petre, M., and Modi, S. (2003). Usability Evaluations of (B2C) e-commerce *Environments, In Proceedings of Workshop 6, Interact 2003, – Exploring the Total Customer Experience: Usability Evaluations of (B2C) e-commerce Environments, INTERACT 2003*, Switzerland.

Monk, A.F. (1998). Experiments are for small questions, not large ones like 'what usability evaluation methods should I use?' *Human –Computer Interaction, 13*, 3, pp. 296-303.

Narayanan, N. H., and Schrimpsher, D. (2000). Extending eye tracking to analyze interactions with multimedia information presentations. *In Proceedings of People & Computers XIV Fourteenth Annual Human Computer Interaction Conference,* Springer-Verlag, pp. 271-286.

National Cancer Institute (NCI). (2002). Research based design and usability guidelines. Available at http://usability.gov/guidelines/ last visited 1st June, 2004.

Nawrocki, L.H. (1972). Alphanumeric versus graphic displays in problem-solving task (Tech. Research Note 227) Arlington, VA: U.S. Army behaviour and systems research laboratory.

Nevalainen S., and Sajaniemi J. (2004) Comparison of Three Eye Tracking Devices in Psychology of Programming Research. *In Proceedings of the 16th Annual Workshop of the Psychology of Programming Interest Group (PPIG 2004).* Institute of Technology Carlow, Ireland, pp.151-158.

Nielsen J. (1999). Alertbox. www.useit.com. Last visited 9th June 2003.

Nielsen J. (2003). Alertbox. www.useit.com. Last visited 10th Nov. 2004.
Nielsen J. (2005). Alertbox. www.useit.com. Last visited 9th June 2003.

Nielsen, J. (2001). *Designing Web Usability: The Practice of Simplicity.* New Riders Publishing, USA.
Nielsen, J., Molich, R., Snyder, S., and Farrell, C. (2000). *E-commerce User Experience*. Nielsen Norman Group, Fremont, CA.
Nielsen, J. and Tahir, M. (2002). *Homepage Usability 50 Web sites Deconstructed.* D. Dwyer, USA.
Nielsen, J., and Mack, R.L. (1994). *Usability Inspection Methods*. John Wiley & Sons, Inc., NY
Nielsen, J., and Phillips, V.L., (1993) Estimating the relative usability of two interfaces: heuristic, formal and empirical methods compared. *In Proceedings of INTERCHI 93,* ACM Press, pp. 214- 221.
Noton, D., and Stark, L. W. (1971) Scan paths in saccadic eye movements while viewing and recognizing patterns. *Vision research, 11*, pp. 929-942.
O'Regan, J. K., Rensink, R. A., and Clark, J. J. (1999). Change-blindness as a result of "mudsplashes". *Nature, 398.*
Oppenheim, A.N. (1992). Questionnaire design, interviewing and attitude measurement. Pinter Publishers, London, UK.
Outing, S., and Ruel, L. (2004). Eye Track III, The Poynter Institute study available at http://www.poynterextra.org/eyetrack2004/ .Last visited 29th April, 2005.
Page, E. B. (1963). Ordered hypotheses for multiple treatments: A significance test for linear ranks. *Journal of the American Statistical Association, 58*, pp. 216-230.
Pan, B., Hembrooke, H., Gay, G., Granka, L., Feusner, M., and Newman, J. (2004). The determinants of web page viewing behaviour: An eye tracking study. *In S.N. Spencer (Eds.), Proceedings of ETRA, ACM SIGGRAPH* .pp. 147-154.
Patel, M.R. (2003). Examining web design conventions across site types. Technical report, Wichita University. Available at: http://psychology.wichita.edu/surl/usabilitynews/51/guest.htm. Last visited 2nd December, 2005.
Pavitt, **K.** (2001) Can the Large Penrosian Firm cope with the Dynamics of Technology? *SPRU Electronic Working Paper Series No.68*, University of Sussex.
Pirolli, P. and Card, S. (1995). Information foraging in information access environments. *In Proceedings of SIGCHI*, Colorado, USA, pp. 51-58
Pirolli, P. L. (2001). *Information foraging theory: recent models and applications.* UC Berkeley Digital Library Colloquium; 2001 October 10; Berkeley; CA; USA.
Preece, J., Rogers, Y., and Sharp, H. (2002). *Interaction Design, beyond human-computer interaction*, John Wiley & Sons, Inc, New York, NY.
Radach, R., Lemmer, S., Vorstius, C., Heller, D. and Radach, K. (2003). Eye Movements in the Processing of Print Advertisements. *The Minds Eye:*

Cognitive and Applied Aspects of Eye Movement Research. Elsevier Science BV, Amsterdam.

Rayner, K., and Pollatsek, A. (1994). *The psychology of reading*. Englewood Cliffs, Prentice Hall, NJ.

Rayner, K., Carlson, M., and Frazier, L. (1983). The interaction of syntax and semantics during sentence processing: Eye movements in the analysis of semantically biased sentences. *Journal of Verbal Learning and Verbal Behaviour, 22*, (3), pp. 358-374.

Reips, U. D. (2000). The web experiment method: Advantages, disadvantages, and solutions. In Birnbaum, M. J. (Eds). *Psychological Experiments on the Internet*. San Diego: Academic Press, 89, pp. 89-117.

Renshaw, J., A., Finlay. J. E., Tyfa D and Ward R. D. (2003). Designing for visual influence: An eye tracking study of the usability of graphical management information. *In Proceedings of Interact*, Zurich, Switzerland, pp. 144-151.

Richardson, D. C., and Kirkham, N. Z. (2004). Multi-modal events and moving locations: evidence for dynamic spatial indexing in adults and six month olds. *Journal of Experimental Psychology: General*, *133* (1), 46-62

Rigney, J. W., and Lutz, K. A. (1976). Effect of graphic analogies of concepts in chemistry on learning and attitude. *Journal of Educational Psychology, 68*, pp. 305-311.

Russell, M. (2005). *Using eye-tracking data to understand first impressions of a web site. U*sability News, 7.1., technical report Wichita University. Available at http://psychology.wichita.edu/surl/usabilitynews/71/eye_tracking.html last visited, 7th Nov. 2005.

Salvucci, D. D. (1999). Mapping eye movements to cognitive processes (Tech. Rep. No. CMU-CS-99-131). Doctoral Dissertation, Department of Computer Science, Carnegie Mellon University.

Savenye, W.C. and Robinson, R.S. (1996). Qualitative research issues and methods: An introduction for educational technologists. In D.H. Jonassen (Eds.) *Handbook of Research for Educational Communications and Technology*. Simon & Schuster Macmillan, New York, pp. 1171-1195.

Schacter, D. L. (2001). *Forgotten ideas, neglected pioneers: Richard Semon and the story of memory*. Psychology Press, Philadelphia.

Schilling, H. E. H., Rayner, K., and Chumbley, J. I. (1998). Comparing naming, lexical decision, and eye fixation times: Word frequency effects and individual differences. *Memory & Cognition, 26*, pp. 1270-1281.

SMI (SensoriMotoric Instruments). (1999). Reference document.

Somervell, J. and McCrickard, D. S. (2004). Comparing Generic vs Specific Heuristics: Illustrating a New UEM Comparison Technique.*In Proceedings of the Human Factors and Ergonomics Society 48th Annual Meeting (HFES '04)*, New Orleans LA, pp. 2480-2484.

Spool, J.M., Scanlon, T., Schroeder, W., Synder, C. and DeAngelo, T. (1997), Web Site Usability: A Designer's Guide, North Andover, MA, User Interface Engineering.

Spool, J. (1999). Spool Vision: An Eye on User Data. Available at http://www.webword.com/interviews/spool.html, last visited, 12th October, 2004.

Spool, J., Schroeder, W. and Ojakaar, E. (2001). Users continue after category link, User Interface Engineering Newsletter, December, www.uie.com .Last visited 5th Dec. 2005.

Stenfors, I., Moren, J. and Balkenius. (2003). Behavioural strategies in web interaction

Stone, L. G. (1941). Reading reactions for varied types of subject matter: an analytical study of the eye-movements of college freshmen. *Journal of Experimental Education, 10* (1), pp. 64-77.

Stone D., Jarrett C., Woodroffe M.R., and Minocha S. (2005). *User Interface Design and Evaluation.* Morgan Kaufmann, Oxford, UK.

Stoy, E. G. (1930). A preliminary study of ocular attitudes in thinking of spatial relations. *Journal of General Psychology,* pp.379-385.

Tatzlaff, L., and Shwartz, D. R. (1991). The use of guidelines in interface design. *In S.P. Robertson, G.M. Oslen (Eds.), Proceedings of ACM CHI'99*, Conference on Human Factors in Computing Systems, New Orleans Louisiana: ACM, pp. 329-333.

Taylor, S. E. (1971). The dynamic activity of reading: A model of the process. *In EDL Research and Information Bulletin*, McGraw-Hill, New York: p. 9.

Tobii. (2004). Eye tracking device specifications available at http://www.tobii.com/, last visited, 3rd Nov. 2004.

Tullis, T.S. (1981). An evaluation of alphanumeric, graphic, and colour information displays. *Human Factors, 23,* (5), pp. 541-550.

Tzanidou, E., Petre, M., Minocha, S., and Grayson, A. (2005). Combining Eye Tracking and Conventional Techniques for Indications of User-Adaptability. *In Constabile, M. F. and Paterno, F., (Eds). Proceedings of Interact 2005, IFIP TC13 International Conference on Human-Computer Interaction*, LNCS, Springer – Verlag, Italy, pp. 753-766.

Tzanidou, E., Minocha, S., Petre, M., and Grayson, A. (2005a). Revisiting Web Design Guidelines by Exploring Users' Expectations, Preferences and Visual Search Behaviour. *In McEwan, T., Gulliksen, J., and Benyon, D. (Eds), Proceedings of 19th British HCI Group Annual Conference*, LNCS, Springer-Verlag, Edinburgh, UK, pp. 421-438.

Tzanidou, E., Minocha, S., and Petre, M. (2005b). Applying Eye Tracking for Usability Evaluations of e-commerce Sites. *In Proceedings of workshop on 'Commercial Uses of Eye tracking' held at the 19th British HCI Group Annual Conference*, Napier University, and Edinburgh; paper available at URL - http://www.amber-light.co.uk/HCI2005/Eyetracking/index.htm

Tzanidou, E., Minocha, S., Petre, M., and Grayson, A. (2004). "An Examination of e-commerce Homepage Design Guidelines by Measuring Eye Movements". *In Dunican and Green (Eds). In Proceedings of the Psychology of Programming Interest Group*, 16th Annual Workshop, Institute of Technology, Ireland, pp. 133-139.

Van Duyne, D. K., Landay, J. A., and Hong, J., I. (2003). *The design of sites: patterns, principles and processes for crafting a customer-centred web experience*. Addison-Wesley, Boston, USA.

Vanderdonckt, J., Limbourg, Q., and Florins, F. (2003). Deriving the Navigational Structure of a User Interface, *In M. Rauterberg, M. Menozzi, J. Wesson*

(Eds.), IOS Press Proceedings of Interact2003, 9th IFIP TC 13 Int. Conf. on Human-Computer, Zurich, Switzerland, pp. 455-462.

Walker, R. Y. (1933). Eye-movements of good readers. *In Proceedings of the Iowa Academy of Science.*

Wickens, C. D., Alexander, A. L., Horrey, W. J., Nunes, A., and Hardy, T. J. (2004). Traffic and flight guidance depiction on a synthetic vision system display: The effects of clutter on performance and visual attention allocation. *In Proceedings of the 48 the Annual Meeting of the Human Factors and Ergonomics Society,* Santa Monica, CA, pp. 218-222.

Williams, K.W. (1999). Navigation display design: Displaying nearest airport information. *Inter-national Journal of Aviation Psychology, 9*, (1), pp.91-106

Yarbus, A. L. (1967). *Eye Movements and Vision*. Plenum press, New York.

Zambarbieri, D. (2005). Commercial uses of eye tracking**.** Workshop of 19th British HCI Group Annual Conference**,** Napier University, Edinburgh**,** paper available at http://www.amberlight.co.uk/HCI2005. Last visited, 2nd December, 2005.

Zola, D. (1984). Redundancy and word perception during reading. *Perception and psychophysics, 36*, pp.277-285.

Glossary

A

Accessibility

The extent to which an application can be used by a variety of people, especially those with disabilities.

Accuracy of task responses

The ability of a measurement to match the actual value of the quantity being measured, in this case the participants' response to the task.

ANOVA (Analysis of variance)

A statistical method for making simultaneous comparisons between two or more means; a statistical method that yields values that can be tested to determine whether a significant relation exists between variables

B

Bright pupil

A bright pupil image is seen by the camera when a light source is placed very close to its optical axis. This effect is well known as the red-eye effect from flash photographs.

C

Calibration

The process or procedure for removing the effect of these subject and set-up differences. A calibration procedure usually requires the subject to look at a certain number of predetermined target points. The relation between the raw measured value (for example separation between pupil and corneal reflection image) and the final device output is adjusted manually or automatically by a computer. It is very advantageous for the calibration procedure to be quick and easy for both the equipment operator and the subject.

Chin rest

A support on which the chin is placed in order to restrict head movements.

Cognitive processes

In psychology it is used to refer to the mental processes of an individual, with particular relation to a view that argues that the mind has internal mental states (such as beliefs, desires and intentions) and can be understood in terms of information processing, especially when a lot of abstraction or concretization is involved, or processes such as involving knowledge, expertise or learning for example are at work.

Conventional evaluation techniques

Conforming to established practice or accepted standards; traditional evaluation techniques such as, interviews, questionnaires, user-observation, think-aloud protocol and so forth.

Corneal Reflection

Also called the first *Purkinjie* image, the corneal reflection or corneal reflex, is the reflection of a light source on the outer surface of the cornea.

E

Eye movement-based metric

A term used to define the measurement of information or a process. More specifically it relates to measuring what happens within an interface either on an ad hoc or ongoing basis based on eye movement data.

H

High-back chair

A piece of furniture for sitting, consisting of a seat, and a high back to allow comfortable position and little head movement.

Homepage

The opening or main page of a website, intended chiefly to greet visitors and provide information about the site or its owner.

I

Infra red light

Light with wavelengths from 0.7 micron to about 0.1 millimeter is called infrared light. The band of infrared light is a thousand times wider than that of visible light. All of it is invisible to our eyes.

M

Menu Bar

ISO Definition: Horizontal set of options, usually located at the top of a work area or window, which invoke lower-level pull-down menus or initiate specific actions. (Source: ISO 9241-14 (1997-06-00) ISO/TC 159).

Multicolumn list

A long list broken into multiple side-by-side columns instead of a single long column.

N

Non-parametric

Non-parametric (or distribution-free) inferential statistical methods are mathematical procedures for statistical hypothesis testing which, unlike parametric statistics, make no assumptions about the frequency distributions of the variables being assessed. The most widely used of these methods is probably the chi-square test.

Norm

Usability norms are regulations for the design of the user interface of applications. Norms may be international and national ones (like DIN 66234/8 or ISO 9241/8), which have the character of laws platform-dependent ones, like the Windows Style Guide or the Apple Human Interface Guidelines.

P

Pixel

A pixel is one of the many tiny *dots* that make up the representation of a picture in a computer's memory. Each such information element is not really a dot, nor a square, but an abstract sample. With care, pixels in an image can be reproduced at any size without the appearance of visible dots or squares; but in many contexts, they are reproduced as dots or squares and can be visibly distinct when not fine enough. The intensity of each pixel is variable; in color systems, each pixel has typically three or

four dimensions of variability such and Red, Green and Blue, or Cyan, Magenta, Yellow and Black.

Point of regard

The correlation of the raw eye position to the precise position on the scene, in real time.

Post-cognitive modelling

A computational approach to the exploration and modelling of cognition.

T

Trend test

The trend analysis procedures are designed to help assess whether there is a functional relationship between the IV and the DV. The functional relationship describes the general trend or nature of relationship between the DV and IV.

Appendices

Appendix A

Background Questionnaire
By filling in this questionnaire you are giving your consent for this information to be used. This information will help to support the credibility of the data gathered.

Name: Age: Gender: ☐ M ☐ F

How often do you use the computer to access the Internet?

☐ Once a fortnight or less
☐ Every day
☐ Once a week
☐ Regularly throughout the day
☐ 2 – 5 times a week

Please tick all the boxes that apply to you.)

Do you use the Internet?

A) ☐ To receive and write email?

B) ☐ To buy or sell general products on-line, such as books, CDs, groceries and so forth?

C) ☐ To buy gifts to be delivered to friends or family?

D) ☐ To buy specialist equipment?

E) ☐ As an information source?

F) ☐ To access discussion lists?

G) ☐ As a business medium (e.g. Intranet for document exchange/maintenance)?

H) ☐ For any other reason (please comment below)

List 3 sites which you visit at least once per week

a) b) c)

List 3 sites that you might not visit once per week but have used a lot and know well

a) b) c)

Pre-session questions for Study 2

- Have you ever looked for an 'About us' link? What happened?
- In which situations do you look up company's information, an example? Background, history, what the company does, and so forth. When do you go to a Website for shopping, travel, information or any other purpose?
- Do you look for the background, history, what the company does? What is their main business? What are their primary products / services?
- <…..>answer
- <ask him to elaborate>
- Why did you look up this information?
- Where could you find it?
- Do you remember where this information was placed?
- Do you frequently look up company information?
- What was the link called? Where was it? Could you find it on the home page?

Pre-session questions for Study 3

- Do you buy tickets or book hotels online?
- What kinds of sites do you like and why?
- What happens when you have images on the homepages of travel websites?
- Do you think it helps?
- Do you like it?
- How about having no images?

Pre-session questions for Study 4

- Have you done online shopping before?
- What kinds of things have you bought?
- What is the procedure you go through?
- Are there any design elements that you use when shopping online? By design elements I mean features such as a shopping cart.
- Where on the interface do you expect to see the shopping cart?

Post-session questions for Study 1

- When visiting a homepage what makes you feel it is actually a homepage?
- What do you look at first when you are visiting a homepage?
- Where might you expect to find the first thing you look for? And why?
- What kind of information do you expect to find on a homepage? What should it look like?
- I have printed out the homepages you just saw. Let's have a look and tell me what do you think you looked at on each homepage to find out about the company?
- Was there something missing from these homepages that you would like to have seen?

Post-session questions for Study 2

- Why do you look for the 'about us' link?
- Where do you expect to find the 'about us' link? In which position?
- Where do you think the 'about us' link should be positioned?
- In which position do you find the 'about us' link?
- What sort of information do you expect to find in the 'about us' link?

Post-session questions for Study 3

- Now, let's have a look at the homepages you just saw.
- How easy was it for you to find what the site was about for the first homepage?
- How easy was it to find the telephone number for the other two homepages?
- Here are some other homepages with different presentation format of the information (show them all the modified homepages), what do you think of these homepages? Some have images, some have text, and do you have any preference?
- Do you think you look at images on homepages?

Post-session questions for Study 4

- First of all how did you feel about the task?
- Was it difficult to find the trolley cart?
- Do you have any preference of the three web pages you just saw?
- Any reason for that?
- Where do you expect to find the shopping cart?
- Here I have print outs of more web pages and the ones you just saw. Have a look and tell me what do you think about them? Which ones do you prefer and why?

Consent form

Thank you for participating in our research program. The research team, from FELS: Dr. A. Grayson, and Computing: Liisa Dawson, Dr. S. Minocha, Dr. M. Petre and Katerina Tzanidou, respect your privacy and will take all reasonable steps to make your comments anonymous. We will not use your name or identifying information in any reports of our research. Your participation is voluntary. You will be asked to a) complete a short questionnaire b) perform a task on a series of homepages of e-commerce sites while your eye movements are recorded and c) to answer a few questions about your experience completing the tasks and your views about homepages. We will be audio and video recording our conversation. All the records, (paper, audio video of eye movements), that are collected will be kept until the end of our research and then they would be destroyed. By signing below you are giving us consent to use the data of this session for our research.

Thank you

Name:

Signature:

Date:

Appendix B

Example template for data organisation per participant

Participant No	**Study 2** *Name of study* **H1** *name of web page* **H2** *name of webpage* **H3** *name of webpage*			
Background Qs	**Sex:**	**Age:**	**Frequency**:	**Purpose**:
Pre-session Qs	***Have you done online shopping before?*** ***What kind of things have you bought? What is the procedure you go through?*** ***Are there any design elements that you use when shopping online?*** ***Where on the interface of the homepage do you expect to find it?*** ***Do you have any preference of where to find the shopping cart?***			
Eye tracking session	**Exposure 1**		**Exposure 2**	**Exposure 3**
Time to target fixation				
Gaze duration text				
Gaze duration image				
Gaze duration other				
Sequence of fixations (scan path)				
Accuracy of responses				
Post-session Qs	***First of all how did you feel about the task?*** ***Was it difficult to find the trolley cart?*** ***Do you have any preference of the three web pages you just saw?*** ***Any reason for that?*** ***Where do you expect to find the shopping cart?*** ***Here I have print outs of more web pages and the ones you just saw.*** ***Have a look and tell me what do you think about them?*** ***Which ones do you prefer and why?***			

Participants' characteristics for Study 3

Characteristics	P a r t I c I p a n t s																													
	Text based Condition										Image Low Condition										Image High Condition									
	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24	25	26	27	28	29	30
Gender																														
Male		•	•	•		•		•	•	•	•		•	•		•					•				•	•	•			
Female	•				•		•					•			•		•	•	•	•		•	•	•				•	•	•
Age	33	24	24	38	23	31	28	30	30	27	23	29	31	27	43	27	43	32	25	45	35	35	24	56	27	27	28	24	49	22
Frequency of internet use																														
Regularly throughout the day	•			•		•	•	•	•	•		•	•		•	•		•	•	•	•	•	•		•	•	•	•	•	•
Every day		•	•		•						•			•			•							•						
Purpose of internet use																														
To receive and write e-mail	•	•	•	•	•	•	•	•	•	•	•	•	•	•	•	•	•	•	•	•	•	•	•	•	•	•	•	•	•	•
To buy or sell general products online	•	•	•	•			•		•	•		•	•		•	•	•	•	•	•	•	•	•	•	•	•		•	•	•
To buy gifts to be delivered to friends or family									•		•			•	•	•	•	•	•		•		•			•	•		•	•
To buy specialist equipment								•		•		•	•		•	•	•	•			•		•	•		•		•	•	•
As an information source		•	•		•	•	•	•	•	•	•		•		•	•		•	•	•	•	•	•	•	•	•	•			•
As a business medium	•	•		•	•		•		•	•	•			•	•	•	•	•		•		•	•			•	•			•
To chat	•				•											•	•									•				•
To listen to music	•				•										•															•
To study online																														•
To read newspapers	•														•															

Percentage of time looking at defined look zones for Study 3

Conditions	Look Zones	Task 1	Task 2	Task 3
Text-based	Text (Content)	**51.02%**	**33.48%**	**80.81%**
	Images			
	Navigation	20.39%	25.25%	7.19%
	Promotional text	4.19%	9.1%	11.06%
	Logo	32.39%	17.72%	
	Telephone No	0.41%	12.46%	
	Address	0.41%		
	Unused	1.19%	1.99%	0.94%
Image low	Text (Content)	**34.38%**	21.57%	26.22%
	Image	26.69%	11.3%	**48.87%**
	Navigation	11.4%	19.32%	5.72%
	Promotional Text	6.72%	**23.08%**	8.61%
	Logo	8.14%	1.77%	3.87%
	Telephone No	7.08%	15.30%	4.16%
	Address	2.50%	3.64%	1.90%
	Unused	3.09%	4.02%	0.65%
Image high	Text (Content)	20.01%	21.33%	24.23%
	Image	**44.11%**	**35.89%**	**56.53%**
	Navigation	9.42%	12.47%	6.01%
	Promotional Text	8.56%	9.58%	9.72%
	Logo	2.68%	9.91%	2.49%
	Telephone No	6.19%	8.38%	
	Address	8.42%		
	Unused	0.61%	2.37%	1.02%

Transition Matrix of fixations for Study 3

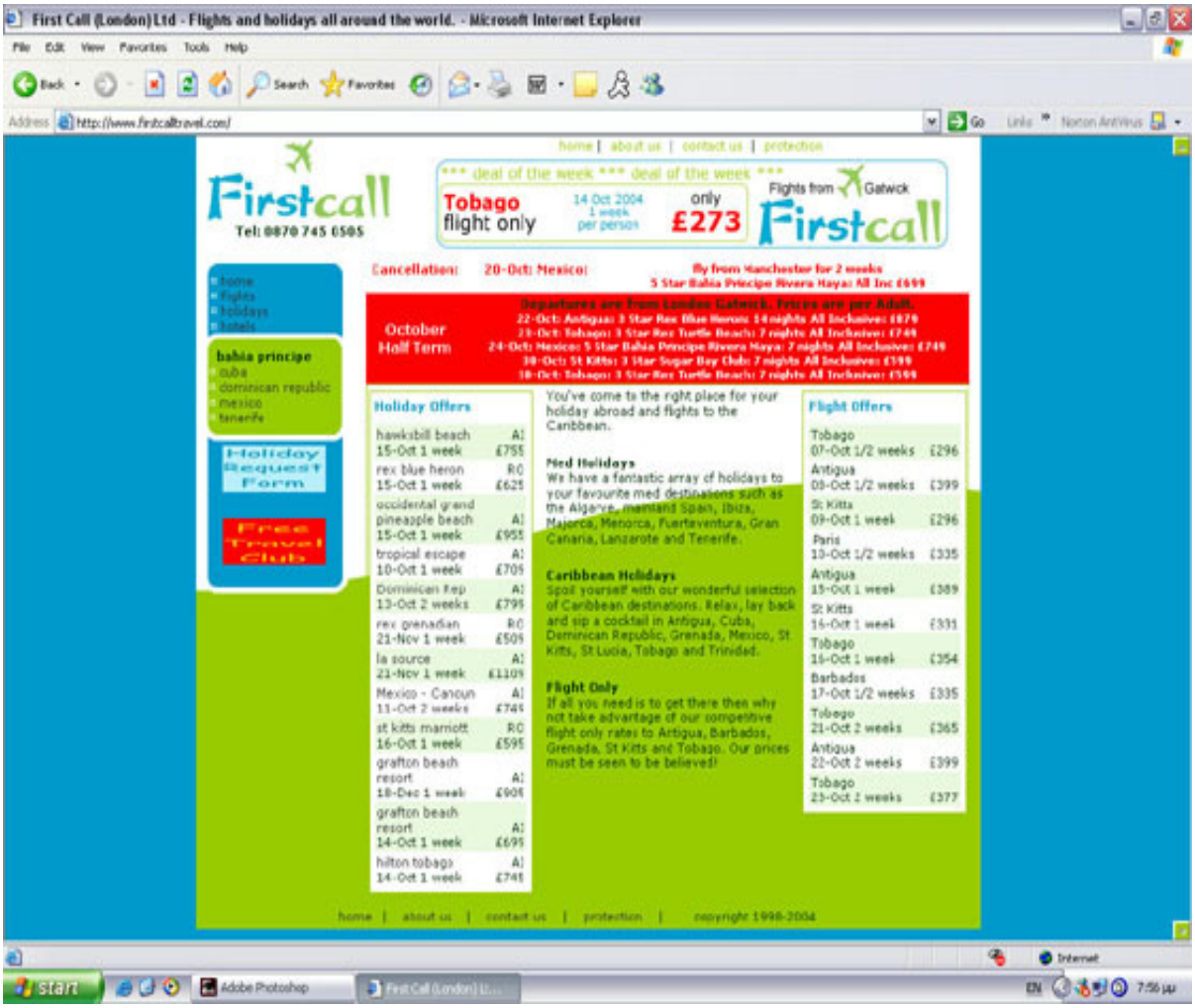
		Task 1						Task 2						Task 3					
Text-based		Same/ relevant text	Different text	Same/ relevant image	Different image	Same/ relevant other	Different other		Same/ relevant text	Different text	Same/ relevant image	Different image	Same/ relevant other	Different other		Same/ relevant text	Different text	Same/ relevant image	Different image	Same/ relevant other	Different other
	T	17.42%	4.54%			4.54%	3.78%	T	12%	6%			6%	4%	T	40%	17.77%			4.44%	
	I							I							I						
	O	9.84%	3.03%			47.72%	9.11%	O	8%				40%	24%	O	13.33%	2.22%			22.24%	
Image Low		Same/ relevant text	Different text	Same/ relevant image	Different image	Same/ relevant other	Different other		Same/ relevant text	Different text	Same/ relevant image	Different image	Same/ relevant other	Different other		Same/ relevant text	Different text		Different image	Same/ relevant other	Different other
	T	20.58%	8.82%	1.17%	0.58%	4.11%	5.29%	T	2.43%	9.75%	2.43%		7.31%	2.43%	T	28.98%	10.14%		8.69%		8.69%
	I	1.76%		2.94%	1.17%	4.11%		I					2.43%		I	7.24%	1.44%		5.79%	1.44%	
	O	4.70%	4.705	5.88%		22.94%	11.25%	O	14.63%		4.87%		43.90%	9.75%	O	1.44%	7.24%	4.34%		5.79%	2.89%
Image High		Same/ relevant text	Different text	Same/ relevant image	Different image	Same/ relevant other	Different other		Same/ relevant text	Different text	Same/ relevant image	Different image	Same/ relevant other	Different other		Same/ relevant text	Different text	Same/ relevant image	Different image	Same/ relevant other	Different other
	T		2.32		6.97%	2.32%	2.32%	T	21.87%	6.25%	3.12%		4.68%	3.12%	T	14.28%	4.76%				
	I	1.16	2.32	4.65%	18.60%	4.65%	3.48%	I		3.12%	1.56%	3.12%	1.56%	3.12%	I	7.14%		21.42%	35.71%	4.76%	2.38%
	O		1.16	8.13%	4.65%	26.74%	10.46%	O	4.68%			4.68%	23.43%	15.62%	O			2.38%	2.38%		4.76%

Detailed data analysis for Study 4

No	Condition	P No	Metric	Exposures		
				Exposure 1	Exposure 2	Exposure 3
1	Text	P3	Initial gaze	B	B	B
			Entry point	Web content	Promotional text	Web content
			Scan path	Processing	Peripheral	Processing
			Web page	Prezzy box	Out of the hat	Beautiful thing
2		P9	Initial gaze	C	B	B
			Entry point	Web content	Promotional text	Top menu bar
			Scan path	Processing	Processing	Focused
			Web page	Prezzy box	Out of the hat	Beautiful things
3		P11	Initial gaze	B	B	B
			Entry point	Web content	Web content	Web content
			Scan path	Focused	Focused	Focused
			Web page	Prezzy box	Beautiful things	Out of the hat
4		P12	Initial gaze	A	A	B
			Entry point	Web content	Web content	Web Content
			Scan path	Peripheral	Processing	Peripheral
			Web page	Out of the hat	Prezzy box	Beautiful things
5		P14	Initial gaze	B	B	B
			Entry point	Web content	Web content	Web content
			Scan path	Focused	Focused	Focused
			Web page	Out of the hat	Prezzy box	Beautiful things
6		P19	Initial gaze			
			Entry point			
			Scan path			
			Web page			
7		P21	Initial gaze	A	A	A
			Entry point	Web content	Web content	Web content
			Scan path	Processing	Processing	Focused
			Web page	Prezzy box	Out of the hat	Beautiful things
8		P22	Initial gaze			
			Entry point			
			Scan path			Peripheral
			Web page			Out of the hat
9		P25	Initial gaze	A	B	A
			Entry point	Web content	Web content	Web content
			Scan path	Focused	Focused	Random
			Web page	Beautiful things	Out of the hat	Prezzy box
10		P26	Initial gaze	B	A	B
			Entry point	Web content	Web content	Web content
			Scan path	Peripheral	Focused	Focused
			Web page	Beautiful things	Out of the hat	Prezzy box

11	**Icon**	P1	Initial gaze	B	A	A
			Entry point	Web content	Web content	Web content
			Scan path	Processing	Focused	Focused
			Web page	Out of the hat	Beautiful things	Prezzy box
12		P4	Initial gaze	A	B	A
			Entry point	Web content	Web content	Promotional text
			Scan path	Focused	Focused	Peripheral
			Web page	Beautiful things	Prezzy box	Out of the hat
13		P6	Initial gaze	B	B	A
			Entry point	Web content	Web content	Web content
			Scan path	Focused	Peripheral	Focused
			Web page	Beautiful things	Prezzy box	Out of the hat
14		P10	Initial gaze	B	B	B
			Entry point	Web content	Web content	Web content
			Scan path	Peripheral	Peripheral	Processing
			Web page	Out of the hat	Prezzy box	Prezzy box
15		P13	Initial gaze	A	A	A
			Entry point	Logo	Promotional text	Logo
			Scan path	Focused	Processing	Peripheral
			Web page	Beautiful things	Prezzy box	Out of the hat
16		P15	Initial gaze	A	B	A
			Entry point	Web content	Web content	Promotional text
			Scan path	Processing	Focused	Focused
			Web page	Prezzy box	Beautiful things	Out of the hat
17		P16	Initial gaze	A	C	B
			Entry point	Web content	Web content	Web content
			Scan path	Focused	Peripheral	Peripheral
			Web page	Beautiful things	Prezzy box	Out of the hat
18		P24	Initial gaze	B	B	B
			Entry point	Web content	Web content	Web content
			Scan path	Focused	Focused	Focused
			Web page	Beautiful things	Out of the hat	Prezzy box
19		P27	Initial gaze	B	B	B
			Entry point	Web content	Web content	Web content
			Scan path	Processing	Focused	Focused
			Web page	Out of the hat	Out of the hat	Prezzy box
20		P28	Initial gaze	A	A	A
			Entry point	Web content	Web content	Web content
			Scan path	Focused	Processing	Processing
			Web page	Prezzy box	Beautiful things	Out of the hat

21		P2	Initial gaze	A	B	B
			Entry point	Top menu bar	Right menu bar	Top menu bar
			Scan path	Focused	Focused	Peripheral
			Web page	Prezzy box	Beautiful things	Out of the hat
22		P5	Initial gaze	A	A	A
			Entry point	Web content	Web content	Promotional text
			Scan path	Focused	Focused	Peripheral
			Web page	Prezzy box	Out of the hat	Beautiful things
23		P7	Initial gaze	B	A	B
			Entry point	Web content	Web content	Web content
			Scan path		Peripheral	Focused
			Web page	Out of the hat	Prezzy box	Beautiful things
24		P8	Initial gaze	B		A
			Entry point	Left menu bar		Promotional text
			Scan path	Focused		Peripheral
			Web page	Prezzy box		Beautiful things
25	**Text and icon**	P17	Initial gaze	A	B	A
			Entry point	Web content	Web content	Promotional text
			Scan path	Focused	Focused	Focused
			Web page	Prezzy box	Beautiful things	Beautiful things
26		P18	Initial gaze		B	A
			Entry point		Web content	Web content
			Scan path		Peripheral	Focused
			Web page		Out of the hat	Beautiful things
27		P20	Initial gaze	B	B	B
			Entry point	Web content	Web content	Web content
			Scan path	Peripheral	Peripheral	peripheral
			Web page	Out of the hat	Prezzy box	Out of the hat
28		P23	Initial gaze	A	A	B
			Entry point	Web content	Web content	Web content
			Scan path	Focused	Peripheral	Peripheral
			Web page	Beautiful things	Prezzy box	Out of the hat
29		P29	Initial gaze	B	A	B
			Entry point	Web content	Web content	Web content
			Scan path	Random	Processing	Focused
			Web page	Prezzy box	Out of the hat	Beautiful things
30		P30	Initial gaze			
			Entry point			
			Scan path			
			Web page			



Text-based condition for www.firstcalltravel.com

Image-low condition for www.firstcalltravel.com

Image-high condition for www.firstcalltravel.com

Text-based condition for www.majortravel.co.uk

Image-low condition for www.majortravel.co.uk

Image-high condition for www.majortravel.co.uk

Text-based condition for www.needahotel.com

Image-low condition for www.needahotel.com

Image-high condition for www.needahotel.com