MSc Human-Computer Interaction with Ergonomics

MSc Project

Web analytics and think aloud studies in web evaluation: understanding user experience.

Elizabeth Atkinson

2007

Tutor: Professor Ann Blandford

optimum·web

usability & accessibility specialists

RedEye

eCRM experts

Abstract

Web analytics and think aloud evaluation methods are commonly used independently to examine users' experiences whilst surfing on websites. This paper explores the value of both methods when used separately through running three independent studies, on the William Hill website. The first phase involved live web analytics data collection and analysis, the second phase involved controlled collection and analysis of web analytics data and the final phase involved running and analysing think aloud evaluation sessions. The results of all three sessions were then analysed to determine the varying types of findings which are gained from each method. The conclusion of this paper presents recommendations of how both web analytics and think aloud evaluation may be used together to overcome limitations of each of the methods and to strengthen insight into user experience. This includes guidelines proposing where both web analytics and think aloud evaluation may be used throughout the User Centred Design process, to help usability practitioners to better understand user experience.

Contents page

Acknowledgements

1 Introduction

Standard Quality Assurance (QA) is not sufficient in ensuring that all visitors to a website are able to complete their objectives through using an interface with good usability, comprising; efficiency, effectiveness and satisfaction (Addwise, 2001). There are therefore a number of further methods which may be used to evaluate the usability of an interactive device, including think aloud, web analytics, cognitive walkthrough, heuristic evaluation and many more. However, there is an understanding within the field that testing using only one technique is insufficient (Nielsen, 2002a). Rosenfeld & Moville (2002) have learned that to get the most value out of user evaluation, integration of multiple research methods is required.

The following paper critically explores specifically web analytics and think aloud evaluation methods. The aim of the study was to determine the effectiveness of each of these methods when used independently. Using findings from this evaluation I devised guidelines on how both web analytics and think aloud may be best applied and possibly combined throughout the User Centred Design (UCD) process to gain a deeper understanding of user experience.

2 Introduction to the context

The study was run in conjunction with optimum.web, a usability and accessibility agency that have a wealth of experience in evaluating websites using the think aloud method. Optimum-web are a subsidiary company of Red Eye International, whose expertise in web analytics facilitated this project.

2.1 Optimum.web and Red Eye International

Optimum.web were founded in 2000, as a pure usability and accessibility agency offering evaluation services based on the principles of Human-Computer Interaction to improve the user experience of websites. The company primarily use expert heuristic evaluation and think aloud usability sessions for web evaluation and have found these methods successful in highlighting usability issues and providing recommendations to improve interfaces. Knowledge and expertise of the think aloud technique was leveraged from optimum.web throughout this project to ensure the data gathering and analysis was realistic to commercially viable evaluations.

Red Eye International are an online behavioural marketing and analysis company who offer tailor made services to help companies understand and maximise their online business (Red Eye International, 2007, promotional pack, RedEye overview page). The method of interest in this paper is RedEye's use of web analytics. Knowledge and technology already in place at RedEye was used to gather and analyse web analytics data used in this study.

In 2006, Red Eye International purchased optimum.web, creating a combined company with a diverse skill set used to better understand online user behaviour. Presently, services of both entities of the company are primarily offered independently.

The following paper explores the overlap between methods used by both optimum.web and Red Eye International to decide how the services may complement each other if combined, using the website of William Hill as an example to test the methods upon.

2.2 William Hill

> *"We invest a great deal in making sure that users are able to navigate*
> *our site as successfully as possible and in this way are driven to*
> *purchase. Working with RedEye, we are able to follow behavioral*
> *patterns closely, with confidence in the data, and most crucially, to*
> *react to customer trends swiftly, providing the best possible solution.*
> *RedEye have added significant, measurable benefit to the business".*

Alex Holt, Internet Marketing Controller, William Hill

(Red Eye International, 2007, promotional pack, Case Studies page).

William Hill is a client of both optimum.web and Red Eye International. Optimum.web use think aloud evaluation to assess the success of the website and have presented recommendations on how designs could be altered to improve user experience based on findings from user tests. Independent work done by RedEye includes evaluation of various designs of the registration pages on the William Hill website, using web analytics to track how many users successfully pass through the registration screens.

In this paper, William Hill is used as a subject of the usability evaluation studies due to access to live web analytics data from this site, available through RedEye technology.

3   Introduction to Usability Evaluation

According to Preece et al. (2002), designers assume that if they and their colleagues can use software and find it attractive, others will too. However, designers cannot be sure that their software is usable and matches what users want, therefore further evaluation is required. Evaluation may be defined as, "the process of systematically collecting data that informs practitioners about what it is like for a particular user or group of users to use a product for a particular task, in a certain type of environment" (Preece et al. 2002). Problems with the design, identified through evaluation represent aspects of the design which, when used in their intended context, cause unexpected results, or confusion for users (Dix et al, 2004).

Evaluation has three main goals: to assess the extent and accessibility of a system's functionality, to assess users' experience of the interaction, and to identify any specific problems with the system (Dix et al, 2004). To meet these goals there are various evaluation techniques. Usability evaluation methods may be broadly categorised into Qualitative and Quantitative data gathering and analysis. Qualitative methods include interviews, observations, think-aloud and some types of questionnaires (Blandford, 2007). Quantitative techniques include other types of questionnaires and web log data.

Furthermore, evaluation techniques may be differentiated by their application by either usability experts, or the involvement of real users. Expert evaluation involves usability practitioners reviewing the design of an interface using their expertise. Wood et al (2003) propose experts' ability to apply cumulative learning and experience from previous evaluations, as what works in web design. However, there is always an element of bias and personal opinions of experts which may devalue the purity of such evaluation. According to Rosenfeld & Morville (2002), the other types of evaluations which involve one user at a time in face-to-face sessions are a central part of the user research process. Preece et al. (2002) also advocate involvement with users, which they state may be highly informative for designers. Benefits of such user participation, discussed by Wood et al. (2003) include flexibility, consideration of non-verbal cues and responses and generally deeper discussions.   Although there are such benefits, there is ongoing controversy in the field of User Centred Design (UCD), questioning how many users should be studied to detect an appropriate number of usability issues (Wixon, 2003). This issue is important in this report, when reviewing the value of think aloud – which generally uses feedback from a limited sample of users and considering this against web analytics data which uses information from the whole user population.

Evaluation techniques may also be either remote from the user or collocated with the user. Such methods may additionally vary in the location of testing, whether it is in the field or in laboratory conditions. The emergence of usability testing in laboratories since the early 1980s is an indicator of the profound shift in attention to users' needs (Shneiderman, 1998). Lab usability testing enables recording of all user actions and key strokes, which provides very detailed and specific feedback on site design which may be used for rigorous analysis (Wood et al. 2003). However, this type of laboratory set-up is very expensive and may affect the number of participants companies are financially able to recruit (Wood et al. 2003). Additionally, where the context of use of a system is critical to the user experience, laboratory testing is not appropriate. As discussed by Preece et al. (2002), laboratory testing does not present an environment which requires users to switch their focus and deal with interruptions which is inevitable in real life when using any interface. This lack of realism, along with the affect of observation is a critical aspect of any testing in controlled conditions. The environment in which data is collected is of particular interest in this paper because most think aloud evaluations of websites are carried out in lab settings, whereas web analytics data is collected from users unobtrusively. This makes comparison of the different types of information which can be sought from such evaluation methods very interesting.

Having introduced the field of usability evaluation, next think aloud and web analytics will be critically explored in detail.

4   Think Aloud

The thinking aloud method has traditionally been used as a psychological research method (Ericsson & Simon, 1984. Cited in, Neilsen, 1993), however it is being used more and more in evaluation of human computer interfaces (Denning et al. 1990. Cited in, Neilsen, 1993). Think aloud usability testing with real participants is one of the most fundamental evaluation methods (Nielsen, 1993). The process involves participants using a system or prototype to complete a predetermined set of tasks, whilst being observed and whilst verbalising their thoughts and comments. The observations and comments are then analysed to define problems in the system which may hinder user experience (Addwise, 2001).

There is a vast amount of literature which critically discusses the method of think aloud. Below, the benefits and drawbacks of this evaluation method are discussed and referred throughout the study in determining how think aloud may be best applied to website evaluation, to gain further insight into user experience.

4.1 Benefits of think aloud

According to Ramey et al. (2006) think aloud is the most important method in the toolkit of usability evaluators because it uncovers more problems than any other measure. Perhaps due to this, think aloud is one of the most popular and frequently applied techniques in testing (Nielsen, 2002a). Although the benefit of a large volume of results is acknowledged in existing literature, this study explores whether some of the limitations of the think aloud method which devalue the results may be overcome using the alternative method of web analytics.

Dix et al (2002) credit the think aloud approach for its simplicity. Along with this, the technique may be described as "quick and dirty", gaining a vast number of results quickly through using simple methodology (Ramey et al. 2006 & Preece et al. 2002). Furthermore, a strength mentioned by Neilsen (1993) is the wealth of qualitative data which may be gained from a fairly small number of users taking part. Although qualitative data is known to be valuable, this study compares findings gained from such data, with information derived from quantitive data, as collected in web analytics. Additionally, the idea of strengthening results by combining findings from methods which independently gather both quantitative (web analytics) and qualitative (think aloud) data will be explored.

The most vital benefit of think aloud is that it provides input to designs from 'actual' users, who are representative of the user population. By verbalizing their thoughts, participants give practitioners an understanding of how they view the computer system, enabling evaluators to identify users' major misconceptions (Neilsen, 1993). The process of think aloud allows practitioners to gain thorough examination of users' behaviour, which can be analysed to divulge the causes of usability problems (Addwise, 2001). Additionally, observational data provides insight into users' affective reactions including; sighs, frowns and scowls which speak of users' dissatisfaction and frustrations (Preece et al. 2002). Although visual clues gained during think aloud are beneficial, this paper compares such findings with results gained from methods which do not include observation, to determine the types of results which may be derived from both obtrusive and unobtrusive methods.

Flexibility of the think aloud approach is also beneficial for usability practitioners. Some facilitators of the method realise that changing conduct during think aloud can affect the kind and amount of data collected. They may therefore manipulate the session to ensure user exploration of certain areas to match the objectives of the test (Ramey et al, 2006). This is a great benefit of the approach when concentrating on certain aspects of an interface which may be suspected to be problematic and to uncover reasoning for the problems.

Despite these positive aspects of the think aloud method, there are some drawbacks which must be considered when examining the value of this technique.

4.2 Drawbacks of think aloud

Ramey et al. (2006) suggest that the validity and reliability of think aloud has been ignored because it is a useful method for convincing developers that problems exist. However, when examining the effectiveness of this method, lack of soundness in the approach is a concern.

Dix et al (2002) suggest the usefulness of think aloud is dependent on the effectiveness of observation and subsequent analysis. This can vary dramatically between usability practitioners, who facilitate think aloud sessions. As stated by Preece et al. (2002), usability testing is strongly controlled by the evaluator, who is in charge. Although this may be positive for the practitioner to ensure goals of the evaluation are met, Hertzum & Jacobsen (2001) believe that due to a combination of vague goal analysis, vague evaluation procedures and vague problem criteria, the think aloud method does not provide evaluators with necessary guidance for performing reliable evaluation. Additionally, Norgaard and Hornbaek (2006) acknowledge concern around the analysis of usability sessions. They suggest that the unstructured format of sessions and incomplete analysis of results may generate findings which do not represent the real user experience. These factors also increase the likelihood of inconsistency in results from different usability practitioners.

Test moderators are also inconsistent in the way they train participants to think aloud, which results in varying amounts of information being divulged by users during a session (Ramey et al. 2006).   The interruptive role of the facilitator in think aloud is also inconsistent which may affect concentration of users, as well as forcing users to spend more or less time on certain aspects of an interface which is misrepresentative of natural use of any device (Ramey et al., 2006, Preece, 1994). Norgaard and Hornbaek (2006) note that moderators may also vary the way they lead tests because they wish for participants to confirm known issues which they may have foreseen or already know are a problem. Throughout think aloud application and analysis of results, moderator interpretation and judgement is required resulting in further differentiation in findings (Hertzum & Jacobsen, 2001). This combination of factors is described by Hertzum & Jacobsen (2001) as the 'Evaluator effect', which adds question to the validity of results from think aloud studies.   Although there is much literature detailing inconsistencies which may occur during think aloud, this study explores where the results of web analytics data, which is

collected unobtrusively, may be applied and combined with the results of think aloud to improve the validity of results from the think aloud method.

Think aloud evaluation involves participants interacting with an interface whilst verbalising their comments, in completely unnatural conditions whilst being observed. This combination of factors reduces the validity of results from this evaluation method. Where users are asked to comment on their experience whilst completing tasks, the increased cognitive load may impact their problem solving behaviour and therefore their task performance (Preece, 1994, Ramey et al, 2002, Preece et al 2002, Neilsen, 1993). Furthermore, thinking aloud may give a false impression of the cause of usability problems if practitioners take users' own "theories" as to what the problem is as gospel (Neilsen, 1993). Moderator judgement and interpretation is therefore required to make sense of user actions, yet this is unique to all usability practitioners and may therefore cause the evaluator effect (Hertzum & Jacobsen, 2001).

The laboratory environment in which most think aloud sessions take place is unnatural and likely to affect user behaviour. Norgaard and Hornbaek (2006) highlight that moderators may negatively effect the way usability sessions are run because they try too hard to meet lab-style scientific standards. This results in the environment being too rigid and not maximising the potential of flexibility during tests to explore all potential usability issues. Participants are not allowed to take telephone calls, talk to colleagues, check email or do any of the tasks which are regularly switched between in normal life (Preece et al. 2002). Observation is also an obtrusive method where users are continually aware that they are being monitored. This may affect their performance and behaviour (McMullen, 2001), a phenomenon commonly referred to as the Hawthorne effect (Preece, 1994). This combination of environmental effects devalue the validity of results from think aloud studies, because the actions of users may be unnatural and not representative of real user behaviour. This study attempts to explore where the findings from web analytics may be combined with those from think aloud to reduce the discussed environmental effects.

Results of think aloud evaluations are limited in what they truly represent in terms of the user population and the interface tested. Running think aloud evaluation on a large sample is time and resource intensive therefore the results of such evaluation usually reflect only a few users (Hong & Landay, 2001). As mentioned previously there is ongoing disagreement around how many users is enough to test upon. Generally between 5-15 users are consulted. Nielsen (1994) suggests that five users are sufficient to find the majority of problems, however a report published by Addwise (2001) states that even when up to fifteen users are consulted, results

do not truly represent the user population.   This suggests that using only a sample of participants reduces the value of findings from think aloud sessions. Additionally, the results are often limited to being representative of only certain aspects of a device or website. Tasks completed in think aloud evaluations are chosen by test facilitators therefore the results concentrate only on a number of predefined scenarios (Addwise, 2001). This limits the value of this method when exploring the vast expanse of a website because areas which may not have been noted by usability practitioners and are therefore ignored.

The concerns affecting the validity of think aloud, mainly caused by the usability practitioner leading and analysing the session as well as the environment in which think aloud takes place, are major flaws in the method. This will be addressed further in this paper, in determining whether unobtrusive data collected in web analytics; which represents real users of a website, may be used in combination with think aloud data to overcome such validity concerns.

5 Web analytics

Web analytics are concerned with collecting, analysing and interpreting web metrics (Weischedel & Huizingh, 2006). The focus is on improving web communication by mining and analysing web usage data to discover interesting metrics and usage patterns (Norguet et al. 2006).   The goal of web analytics is to assist companies in improving the quality of their website – including the usability and performance of the site (Weischedel & Huizingh, 2006). RedEye's data analytics service can help businesses identify aspects of e-business which are working well through reports and then devise recommendations of how they may work better (RedEye, Analytics promotional pack, 2007).

Chi (2002) proposes that poor website usability, results from designers' lack of understanding of how users surf for information. However, correct analysis of web analytics data can present this type of information. Data can be tracked relating to: how long people stay on a website, which areas they visit, where they came from and where they went next (Preece et al. 2007 & Sterne, 2006). This data is generally presented in reports which may be used to examine the efficiency of regularly used aspects of a website, as well as reviewing rarely used facilities to understand why users may avoid such features (Shneiderman, 1998). The usefulness of such web analytics reports in corporate decision making depends on the report viewer; web designers are interested in detailed reports, whereas strategic managers only require summaries which show the number of visitors to each page (Norguet et al. 2006).   As discussed with Andrew Stockwell, Account Director at Red Eye International, reports may be

adjusted to either explore fine details of certain aspects of a website or present broad information on site usage, to meet the requirements of the person using the details.

There is some literature which details the process of collecting and analysing web analytics data, along with critical discussion of the method. Much of this literature details specific software which is used to collect and analyse web analytics data. However, the following section details general benefits and drawbacks of the method on the whole, when used to gain an insight into user experience.

5.1 Benefits of Web analytics

Web analytics data is widely available in large volumes, as it is stored by default in server logs and through page tags (Weischedel & Huizingh, 2006, Preece et al. 2002). The data provides usage information from all users, which has a fairly low cost per use (Addwise, 2001). This ensures that results truly represent the behaviour of the whole user population and are not restricted by financial or resource limitations. Use of this large volume of data is explored later in this study when comparing the value of such vast results against the findings of a fairly limited sample of one-to-one think aloud sessions.

Collection of web analytics data is unobtrusive (Preece et al. 2002). Data is collected outside of the laboratory, while users are engaged in real tasks, in their natural environment, so the results have greater ecological validity because users behave more realistically (Gurdzial et al., 2004 & Addwise, 2001). Furthermore, the validity of data is increased because it is gathered discreetly and therefore the possibility of a Hawthorne effect is eliminated (Gurdzial et al., 2004). Comparing the results of web analytics with observed think aloud sessions is very interesting because it may uncover how the results differ. This may be representative of the effect of observation and the controlled environment in which think aloud sessions take place. Additionally, during think aloud studies a scenario is usually given to participants with specific tasks which they must complete. This restricts the potential pages they may visit to only those decided on worth testing by the usability practitioner prior to the session. These limitations of think aloud sessions are overcome in web analytics data because they are not restricted to pre-determined evaluators concepts and tasks (Addwise, 2001) and therefore results can potentially represent the whole interface. However, often web analytics data does not cover all site traffic because of untagged areas of the site, which results in incomplete available data (Weischedel & Huizingh, 2006).

5.2 Drawbacks of web analytics

The problem with data from server logs is the vast volume that gets generated, which makes it difficult to determine what is useful, and how to display the useful portion in a meaningful way (Guzdial et al., 2004). The data requires extensive pre-processing to provide insight into user behaviour which is costly in time and resource (Spiliopoulou, 2000 & Chi, 2002). The metrics produced by web analytics contain little semantics and are too detailed to be exploited by organisations, therefore require expertise to summarise data so that it can be used to make decisions (Norguet et al. 2006). Powerful tools may be used to explore and analyse the data (Preece et al. 2002), but these are expensive to create and may generalise certain findings. To overcome such generalisations expert consultants may analyse the data to generate more specific results. However, as found in the Web Analytics survey by Peterson's (2007), over 50% of respondents described web analytics as difficult or extremely difficult – therefore expertise in this area is not easy to attain.

As found in Peterson's (2007) survey, the majority of people coming into contact with web analytics data from their website do not understand what the data means. This is due to its complexity as mentioned by a number of authors (Spiliopoulou 2000, Chi 2002, Preece et al. 2002). This facet of the data makes it more difficult to use to base strategic decisions upon because organisations are unable to gain a clear understanding of what the data represents in terms of use of the website, and how this may be used in decision making. When questioned in Peterson's (2007) survey, 65% of respondents indicated that a top organisational concern was the ability to use web analytics data for decision-making purposes. This suggests that web analytics has limited value for strategic purposes.

Since users are unaware of the collection of web log data, there are ethical considerations which must be taken account of when analysing such data (Preece et al. 2002). Logging may be well intentioned, but users' rights to privacy deserve to be protected (Shneiderman, 1998). Shneiderman (1998) also states that links to specific user names should not be collected unless necessary and suggest that managers must inform users of what is being monitored and how the information will be used. This adds to the amount of processing required by analysts to ensure data can be understood by all people reading the web analytics report.

As found by Peterson (2007) in the Web Analytics survey, a general problem with web analytics is that an incomplete picture of website usage is provided. Often the data does not cover all site traffic because of untagged areas of the site which results in incomplete available data (Weischedel & Huizingh, 2006). Where gaps appear in web analytics data, analysts are

required to make assumptions to create a fuller picture which limits the accuracy of the conclusions drawn from the data. This study attempts to find out if the results of think aloud may be combined with the data from web analytics to reduce the requirement for such interpretation.

Web analytics data is not only limited in the breadth of the website covered but also in the information it presents about user actions. Analytics data displays the 'when' and 'what' of web visits, but is of limited value when answering 'how' and 'why' questions about customers' site use (Weischedel & Huizingh, 2006). As described by Hong & Landay, (2001), web analytics data resembles "footsteps in the forest" – you know where someone went but you have no idea what they were trying to do and whether they were successful. Addwise (2001), concur that no current web analytics tool offers indication into reasons behind users' actions. This is one of the main flaws addressed in this project, determining whether think aloud evaluation may be used in combination with web analytics to provide more information of why users follow the paths they choose.

6 Methodology of investigation

6.1 Overview

The methodology used in this study consisted of three main stages of data gathering and analysis. The initial stage involved live web analytics data, the second phase used controlled web analytics data and the final stage involved think aloud evaluation. Live web analytics data was gathered and analysed first to provide me with experience of using real analytics data and to determine how much can be inferred from such information with no awareness of user objectives. This was also the primary phase because it enabled me to define tasks to be completed in the proceeding stages of evaluation, based on findings from live data.

The next phase of data gathering and analysis was controlled web analytics. This involved monitoring use of the William Hill website by known users recruited by myself, who completed tasks also set by myself. This stage was completed prior to think aloud, because it was important that findings from think aloud – which were likely to be more informative, did not influence the analysis of web analytics data. Having explored known flaws of web analytics in the previous chapter of this study, it is clear that lack of knowledge about what users are thinking is a problem when analysing the data. However, this information is divulged in think aloud, therefore if think aloud evaluation had taken place prior to web analytics analysis this valuable information may have been transferred, damaging the purity of the study.

The final phase of data gathering and analysis was think aloud evaluation, which was script based and used the same tasks as completed in the controlled web analytics study, to ensure comparable data. As described earlier, this was completed following the earlier phases because it was likely to provide the most detailed information of usability issues.

Following all data gathering and analysis, a final phase of drawing all findings together to compare the results of each of the tests was completed to present results and guidelines of how each evaluation method may be best used throughout the User Centred Design process.

The study therefore involves a number of phases of gathering and analysing of both qualitative and quantitative data. Prior to these main stages however, I was introduced to the process of gathering and analysing web analytics by Mr Andrew Stockwell, Account Director at Red Eye International.

6.1.1 Induction to web analytics

Having had very little experience using web analytics, I was introduced to the RedEye Reporting Interface by Mr Stockwell. This interface has been created uniquely for Red Eye International, and is used by both consultants and clients to create reports from web analytics data which is collected by RedEye tagging technology.

Reports are generated through selecting criteria of which aspects of user pathways are to be included in the report, as shown in figure 1. Refining selections of criteria allow the person creating the report to drill down into certain parts of the data, which can reflect detailed user pathways within specific sections of a website.

Figure 1 shows the RedEye reporting interface, when being used to create a report based on the William Hill website. The criteria for this report has been selected to show all user pathways which contain the following three separate pages: i) Home Page ii) Log in Success and iii) Bet Slip. The meaning of 'Home Page' and 'Log in Success' is self explanatory, however through exploration of the William Hill website it is clear that the Bet Slip is where users of the website can view all their open bets.

Figure 1: RedEye Interface, URL selection screen for the William Hill website.
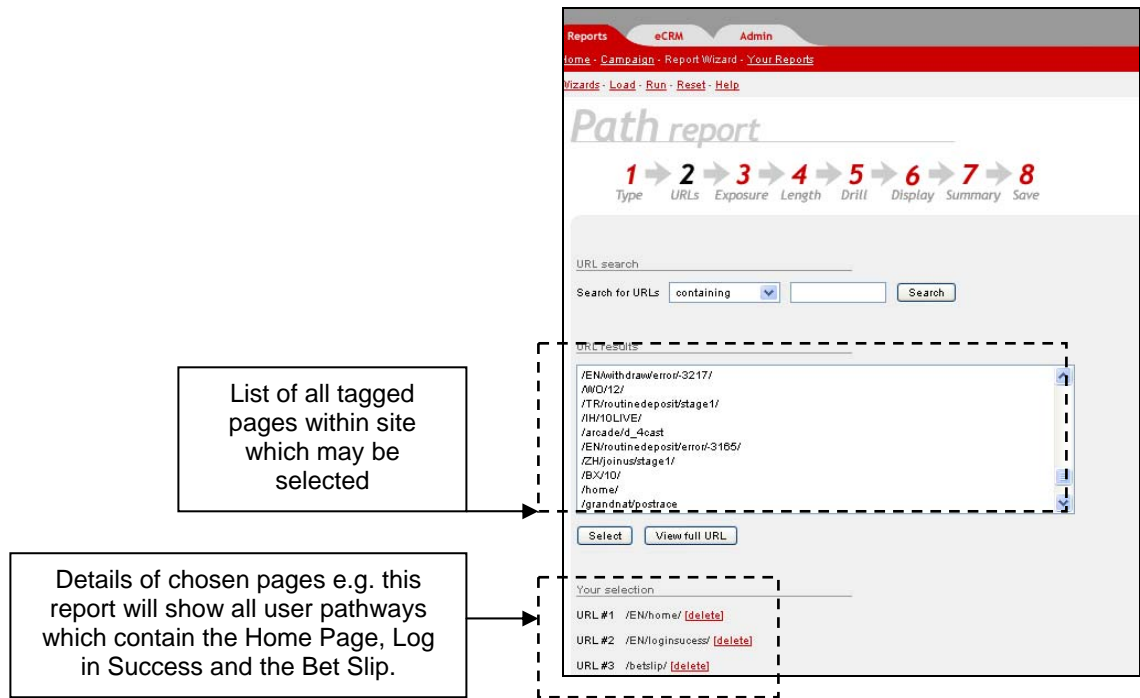
Additionally, selections may be made to present only data collected over a specific period of time, which may be particularly helpful if reviewing how many views a certain item had when released on a certain date.

Having completed all criteria selections the report is created in either HTML or CSV format. An example of an HTML report is shown in Figure 2. The report details the number of visitors, visits, new visits, repeat visits as well as the path time in seconds for all users whose pathways matches the criteria fixed in the report set up. In this study only the number of 'Visitors' was concentrated on. This report may then be used by analysts to draw conclusions regarding site activity.

Figure 2: Example of a RedEye web analytics path report
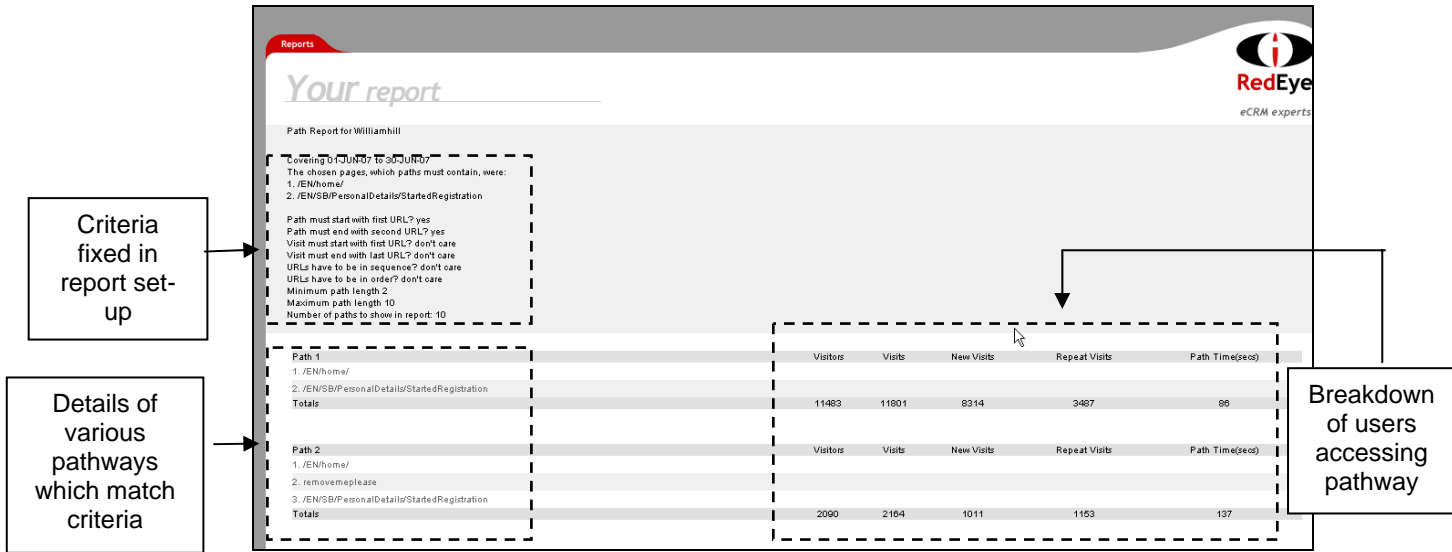
The induction to the RedEye Interface involved Mr Stockwell showing the types of reports which may be drawn from the interface, as well as explaining how the data may be analysed to further understand user pathways. This information was critical to my data gathering and analysis later in the project, to ensure the analysis was realistic to that of commercial web analytics evaluation.

Following this induction, I was able to commence the data gathering phase of the project.

6.2 Data gathering

6.2.1 Ethical clearance

Prior to all testing, ethical clearance was granted from the UCL Research Ethics Committee. In line with this, all participants consented to taking part in the tests and were informed that they were able to withdraw at any point. See Appendix 1 for ethical consent forms.

6.2.2 Participants

The live web analytics data used in this study represented the pathways of the real user population of the William Hill website whose personal details I did not have access to. This is realistic to true web analytics data, where users are often anonymous. Due to limited resources I personally recruited the participants for the controlled web analytics and think aloud phases of this study. To retain validity of the results I attempted to recruit a balanced sample, including users who represented a wide age range and varied technical experience. The

independent web analytics and think aloud tests were between subjects, using completely unique users whose details are given in the relevant sections of this paper.

6.2.3 Live web analytic data

The first phase of this study involved reviewing live web analytics data from the existing William Hill site. I began by exploring all data available from the RedEye Interface to gain overview information about different areas of the William Hill site which were visited. From this, I could then see which areas were most popular and any areas of the website where user pathways were unusual. Having completed independent data gathering I then discussed the data with Mr Stockwell, who has experience in working with William Hill and was aware of issues which the website owners are most interested in exploring. During this discussion, I was informed of the significance of the registration process of the website. Using this client information I was then able to drill down on data showing user pathways during the registration process by selecting criteria in the interface to report only user pathways which included access to pages within the registration process.

This phase provided me with experience of analysing real web analytics data, helping to inform me of the benefits and drawbacks of using such data to evaluate user experience. When working independently, I was able to draw conclusions purely based on the number of visitors accessing certain areas of the site and the types of pages they faced – including error messages. With information provided by the William Hill client I was then able to gain deeper insight into what they realised were areas where users had difficulty. This information was very useful to ensure that the results of the review were beneficial to the client organisation. Although through purist web analytics analysis I was able to present findings of web traffic, it is also useful to have knowledge of which aspects of user pathways clients are most interested in, and are important to site success.

The most prominent problematic areas of the site found using the results of this review, within the registration process were then used to base tasks for users to complete in the proceeding stages of the study; controlled web analytics and think aloud. To ensure thorough exploration of the site, and a more realistic scenario during the proceeding tests, the tasks involved users placing bets on sporting events as well as registering with the William Hill site.

6.2.4 Controlled web analytics

The controlled web analytics test involved thirteen participants using the William Hill website. Participants were asked to complete tasks set by myself in their own time and environment.

Users were emailed with instructions (see Appendix 2) and an ethical consent form (see Appendix 1) from myself and given a deadline, by which time they must have completed the tasks. Following previous agreement to participate by twenty five users, I emailed them all with instructions for the test. However, unfortunately only thirteen users completed the tasks. Additionally, as shown in table 1, the pathways of two users could not be found in the logs.

Users' first task was to register with the William Hill website. This involved them using a pre-set username provided in the instructions. As can be seen in the instructions in Appendix 2, it was necessary for me to inform users that they were not to proceed with the tasks until they were shown a confirmation of their Account Number. Although this may add bias to the results, because I provided users with extra information of how to know when registration was complete, this was necessary to ensure I was able to pull out individual user's sessions, which are identified by Account Numbers in the web log data. If users did not manage to get this page their data would not be accessible to me and therefore there would be no point in them continuing with the tasks.

Following registration, users were asked to place bets on two events, which they had to locate within the William Hill website. The first event was the 2009 Ashes cricket tournament and the second was the 37th golf Ryder cup. Having placed bets, users were asked to log out of the website to close the test. RedEye technology tracked the pathways of these users during task completion. Following this, participants informed me of the exact time and date that they completed the tasks, as well as their Account Number. When all participants had informed me of these details, Andrew Stockwell pulled the data of pathways taken by these users from RedEye web data logs.

Table 1: Controlled web analytics, participant details

Grey text shows details of users whose data was not found in logs.

Account Number	Sex	Age	Nationality	Use of Internet	Use of online betting sites	Use of offline betting methods
22592O	Male	36	British	Frequent	Occasional	Occasional
181530	Female	28	Canadian	Infrequent	Never	Never
14849O	Male	25	British	Frequent	Occasional	Occasional
14030O	Male	69	British	Infrequent	Never	Occasional
22562O	Female	32	British	Infrequent	Never	Never
12834O	Female	24	British	Frequent	Never	Occasional
17140O	Male	23	British	Frequent	Never	Occasional
216680	Female	22	Chinese	Frequent	Never	Never
21637O	Female	22	Greek	Frequent	Never	Never
23391O	Female	23	British	Infrequent	Never	Occasional
281260	Female	23	British	Frequent	Occasional	Never
282750	Female	26	British	Frequent	Never	Never
282730	Female	26	British	Infrequent	Never	Never

6.2.5 Think aloud

The think aloud evaluation involved eight participants using the William Hill website to complete tasks set by myself in a usability laboratory set up. A task script was created, detailing the tasks to be completed by users (see Appendix 3). These tasks were consistent with the tasks done in the controlled web analytics phase, to ensure comparable data. In addition to task information, the script also included trigger questions to remind the moderator of specific aspects of various pages of the website which may be explored to uncover any usability issues.

Users were requested to use the think aloud protocol when attempting to complete tasks, whilst being observed by myself acting as the moderator of the session. Notes were made of the routes taken, and comments made by users to be referred to in later analysis. Additionally, Camtasia 3 recording software was used to record sessions which was reviewed in analysis where notes were not sufficient. The details of all five participants of the think aloud sessions are shown in table 2.

Table 2: Think aloud, participant details

User Number	Sex	Age	Nationality	Use of Internet	Use of online betting sites	Use of offline betting methods
1	Male	25	British	Frequent	Occasional	Occasional
2	Female	63	British	Infrequent	Never	Occasional
3	Female	29	British	Frequent	Occasional	Never
4	Male	25	British	Frequent	Occasional	Occasional
5	Female	24	British	Frequent	Never	Never

6.3 Data analysis

6.3.1 Live web analytics data analysis

Data from the live William Hill site was analysed using the RedEye Reporting Interface which allowed me to set parameters within the data to create web analytics reports. Initially I set fairly open criteria to gain overview information of user pathways throughout the William Hill site. Analysis of this information included interrogation of the reports to explore any trends, or anomalous figures within the data which suggested unusual or incomplete pathways taken by users. Additionally, I was able to view where users were faced with error messages when using the website. Where this arose, I then explored the surrounding pages of error messages, to find out at what point errors occurred and why they may have occurred. As well as errors, I was able to explore particular areas of the website by selecting from the RedEye Reporting Interface specific start or end points or certain pages accessed during user journeys. Throughout this phase it was critical for me to make assumptions of users' objectives throughout using the site and whether these goals were met or hindered during their session.

Having found out through Andrew Stockwell that William Hill clients were most concerned about user behaviour during site registration I was able to analyse user journeys which included these pages in more detail. This information was useful to explore more specific areas of the data, which made it much easier to draw conclusions. As mentioned previously, the RedEye interface allows for a vast range of information to be presented, however it is difficult to draw conclusions efficiently without any guide of which sections of the data to explore.

6.3.2 Controlled web analytics data analysis

Unfortunately, of the thirteen participants who completed the controlled web analytics test – two user details were not shown in the RedEye web logs, so only eleven user pathways were analysed.

Analysis of this data began by removing all extraneous information provided in the logs which would not be helpful during my exploration of user pathways. This included details of the I.P address and the browser used by participants. Following this, I reviewed the pathways of all users initially by evaluating the number of users who successfully met the objectives, set in the instructions given to participants. This helped to establish any high level issues with completing the tasks. Having noted overview details of pathways, I then explored further to find out if there were any unusual pages shown within the pathways. This included error messages, unknown sporting event pages, repeated viewing of the same page and any visits to the Help pages of the site during the visit. Where any of these anomalies occurred I reviewed the William Hill website in an attempt to establish what the users may have been attempting to do when faced with these pages. There were some codes used in the log data which I could not make sense of, due to my inexperience in analysing web analytics. It was therefore necessary for Mr Stockwell of RedEye, to inform me of the meaning of some of the codes used in the log data. Having established all the paths and possible issues users had whilst completing the tasks, I prioritised the findings to be documented in the results section. This prioritization involved reviewing the number of users affected by each issue, as well as how influential it was in completing the tasks set.

6.3.3 Think aloud analysis

Analysis of the think aloud sessions involved recalling observations and reviewing both transcripts and Camtasia 3 recordings from the sessions. This combination of sources of qualitative information gained provided a vast amount of information to be explored to uncover usability issues. Transcripts from the sessions detailed all aspects of the site and experiences commented on by users, so this was the most useful material in the analysis. Initially I looked for trends in the data to find any critical issues which repeatedly occurred during user sessions, therefore affected multiple users. Additionally, using my HCI expertise gained through previous experience in running and analysing think aloud sessions, I was able to uncover issues which were inferred by user actions, even if they were not directly vocalised.

Having prioritised the most critical issues based on the number of users they affected, I explored the reasoning for the problems. This involved reviewing what users said as they

faced the issue, along with their actions. Using this combination of evidence ensured that I did not solely rely on false impressions which may arise from thinking aloud, nor take participant's "own theories" as gospel as described by Neilsen (1993). Having understood why each user had their unique difficulties I looked for trends in this reasoning, to uncover the most critical usability issues which affected multiple users for the same reasons. Details of these results are described in the think aloud results section of this paper.

7   Results

The results section includes a detailed breakdown of the findings gained from all three studies conducted; live web analytics, controlled web analytics and think aloud. Following the results from each study there is a short reflection on the methodology used.

7.1   Live web analytics data

The following section details the findings from web analytics data, gained when using live data collected from the William Hill website. Overall, the data was beneficial to gain top level understanding of the pathways taken by users in their own environment whilst behaving naturally on the website. However, as detailed in the proceeding sections it is difficult to draw conclusions from such broad data with no further user consultation.

7.1.1   Top level website usage information

When analysing the live data from the William Hill website, without guidance of any specific areas of data to explore, it was difficult to analyse the data efficiently. It is tempting to continually alter criteria to focus on various aspects of the site, however without knowing the site in detail and being unable to infer why users may have visited certain pages it is challenging to draw conclusions. This is in line with published literature by Spiliopoulou (2000) and Chi (2002) which proposes that without extensive pre processing, it is challenging to gain useful results from web analytics data.

The vast range of criteria which may be chosen from the RedEye Reporting Interface allows for wide exploration of the various pathways taken throughout the whole website. However, this can make finding information very hit or miss and it is time consuming for the analyst. Although this type of analysis is useful for gaining a top-line overview of user pathways through the website, it is difficult to uncover where issues are without any prior indication of where to start looking.

Having been informed in a meeting with Mr Stockwell, that the site owners of William Hill had concerns around user experience during site registration, I focussed on data representing this process in more detail. This information of a specific area of interest was useful to ensure I explored parts of the data which would help to examine the success of the website, rather than very general information about use of the website which was difficult to draw conclusions from.

The following sections of this paper provide examples of problems found on the William Hill website which were highlighted in the live web analytics data. Also there is discussion of how these issues may be used by usability practitioners to enhance their understanding of user experience.

7.1.2 Highlighting issues

7.1.2.1 Log in failure

In reviewing the live analytics data it is clear that many users have difficulty logging into the William Hill site, facing login failure error messages. The data in figures 3 and 4 show the two most taken paths by users who successfully log in to the site.

Figure 3: Number of users successfully logging in from the Home page.

1. /EN/home/	
2. /EN/loginsucess/	
TOTAL VISITORS	**126,611**

Figure 4: Number of users facing an error message before successfully logging in from the Home page.

1. /EN/home/	
2. /loginfail/	
3. /EN/loginsucess/	
TOTAL VISITORS	**23,143**

From the data in figures 3 and 4 it is clear that, for every 100 people who successfully accessed the site, 18 people faced an error message prior to logging in, which is a concern for the site owners. Web analytics are valuable in highlighting this issue to site owners, however the pure data does not give any indication of why users have such problems.

As can be seen in figure 5 which shows the log-in section of the William Hill home page, user issues could occur due to the visibility of labels, wording of headings or issues with

remembering and entering correct user details. However, web analytics data only shows that error messages did occur and it is not clear from the data where the issues in the interface stemmed from.

Figure 5: Log in section of William Hill website

7.1.2.2   Errors during registration

Further error messages occurred on user pathways during the registration process. When people began the registration process, a number of users faced error messages regarding their security details.

Figure 6 and 7 show the two most taken pathways by users who accessed the Started Registration page of William Hill, which is the first page of the registration process. These pathways represent data recorded between the dates of 01/06/2007-30/06/2007.

Figure 6: The most taken pathway which included the Started Registration page

1. /EN/home/	
2. /EN/SB/PersonalDetails/StartedRegistration	
3. /EN/SB/PersonalDetails/ManualAddress	
4. /EN/SB/FindUKAddress/SelectUKAddress	
5. /EN/SB/SelectUKAddress/ConfirmUKAddress	
6. /EN/SB/ConfirmUKAddress/SecurityDetails	
7. /EN/SB/SecurityDetails/error	
8. /EN/SB/SecurityDetails/error	
TOTAL VISITORS	**4249**

27

Figure 7: The second most taken pathway which included the Started Registration page

1. /openAccount/confirm/	
2. /EN/SB/PersonalDetails/StartedRegistration	
3. /EN/SB/PersonalDetails/ManualAddress	
4. /EN/SB/SB/SecurityDetails	
5. /EN/SB/SecurityDetails/AccountConfirmation	
TOTAL VISITORS	**2874**

As can be seen from the data in figures 6 and figure 7 for every 67 users who registered successfully with the William Hill site without facing any errors, 100 users did face error messages. This is a huge number of users facing an error which will have greatly affected users' experience of the site. For William Hill, the fact that users are facing an error message is critical to successful use of the site and the problem must be addressed. However, it is not clear from the pure data how to go about rectifying the problem.

As shown in figure 6, in row six - the error faced by users occurred, after their access the page "/EN/SB/ConfirmUKAddress/SecurityDetails". It is not clear from this data the reasoning for this error message, however having reviewed this page of the registration process, there are potentially at least six reasons for an error message, as shown in figure 8.

Figure 8: User registration screen where users faced many errors.

Figure 8 was purposefully generated to highlight the potential errors which users may be faced with when attempting to complete this page of registration. The errors highlighted in red include, users failing to enter a suitable user name, failing to enter a password which matches the criteria, failing to select a security question, failing to enter the answer to their security

question, failing to enter a daily deposit and/or failing to accept the terms and conditions of the William Hill site.

Having discussed this error with Mr Stockwell, he informed me that RedEye analysts can access detailed error logs which provide a specific breakdown of errors faced by users, however these are not shown in the path reports which I had access to. This suggests that the exact field with which users had difficulty could be located in more detailed web analytics data, however even this does not provide information of why users faced the error.

To use this data to understand in more detail the problems users faced, it is necessary to review this page and potentially use expertise and experience in usability of forms to establish why errors may occur. Additionally, observation of users who came across these errors would provide clues of why the error occurred.

7.1.2.3 Incomplete registration

The result of errors faced during the registration process not only hinders user experience when trying to register but as can be seen from the data in figure 6, a number of users abandoned their transaction and left the website after being faced with an error message. This can be read in the table, because the last row of the table contains an error message, yet if users had persevered with registration, further pages would be listed in additional rows. As mentioned in the previous section the number of users who gave up on registration having faced an error is dramatic, at 100 users for every 67 users who successfully registered.

For William Hill, this means that such a large number of users were not completing registration and therefore will not have placed bets on the site and most importantly will not have spent money on the site. Knowledge of such a vast loss of this type of 'almost registered customers' is vital for William Hill because these users did have the intention to register with the site, therefore are interested in the company's offerings – however failure messages during registration negatively affected their experience enough for them to leave the site. This is obviously a great concern for the success of the company.

This fact is in line with literature published by Rimmer et al. (2000) who studied the effect of web error messages. They found that users were irritated by error messages and seldom made the effort to recover from the situation unless it was very important to them. In response to this, Rimmer et al. (2000) recommend using better-informed error messages to make the process of recovering from error easier and therefore result in more successful interaction.

This suggests that William Hill should review the error messages provided to further aid users in recovery.

Information provided by web analytics is useful to highlight problems which are occurring and may hinder user experience, however so far the pure data did not give any indication of why users faced errors.

As well as errors, the data shows that a large number of users also ended their pathway and therefore left the site –having only started the first page of registration, as shown in figure 9.

Figure 9: Number of visitors leaving site on first page of registration.

1. /EN/home/	
2. /EN/SB/PersonalDetails/StartedRegistration	
TOTAL VISITORS	**1620**

This implies that there is a problem on the first registration screen, causing users to leave the site. However this data may alternatively represent users who mistakenly click on the Open Account button from the home page. Although the results of the web analytics are useful to highlight this issue to site owners, to find out the actual reasoning for the occurrence it is important to speak to real users of the site.

Having realised the benefit of web analytics in highlighting problems users faced as well as unusual drop off points, the following section describes the use of web analytics in presenting data regarding how the site is used by real customers.

7.1.3 Realistic user behaviour

Web analytics data is collected when users are using a website, with no awareness of their pathways being logged by tracking software. This is a great asset of web analytics, which is impossible to create in think aloud evaluation. This factor is useful when using analytics to rate the popularity of pages. If it is highlighted in analytics that certain pages are visited more frequently than others, designs may be changed to ease access to the most popular pages.

Figure 10 shows a screen shot of the 'My Account' section of the William Hill website. Quantitative values taken from the live web analytics data show the number of visitors who accessed specific parts of this page prior to logging out, on the 24th June 2007.

Figure 10: My Account page with live visitor usage data.
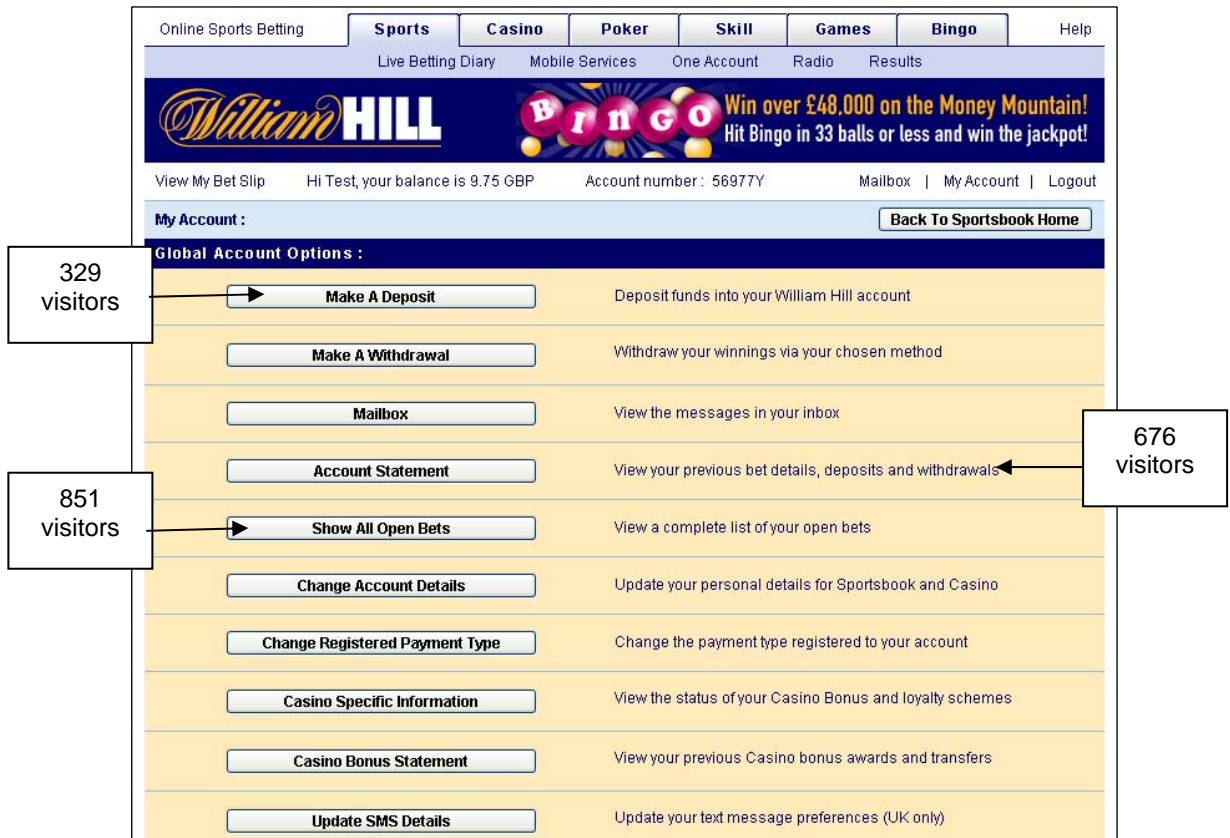
When exploring the pathways of users who viewed the My Account page, it was clear that most users (851 users) went to View All Open Bets first. However, the list shows View All Open Bets as the fifth item in the list of options. This type of information shown in the analytics may be useful to inform simple design changes of the pages to improve user experience by easing access to the most used items. However, although this may ease user experience it is likely that site owners have reasoning for the positions of links on this page. For example, site owners may try to heighten awareness of the Make a Deposit link, by positioning it at the top of the page to increase the likelihood of users spending money on the site. It is therefore important to combine the findings from web analytics with knowledge from site owners before using results to influence design decisions.

Further data from the William Hill website shows that a number of users go from the My Account page of the website to the Show All Open Bets page and then move through Account Statement area. As can be seen in figure 11, this is an indirect route to the Account Statement – which actually has a separate link.

Figure 11: My Account page
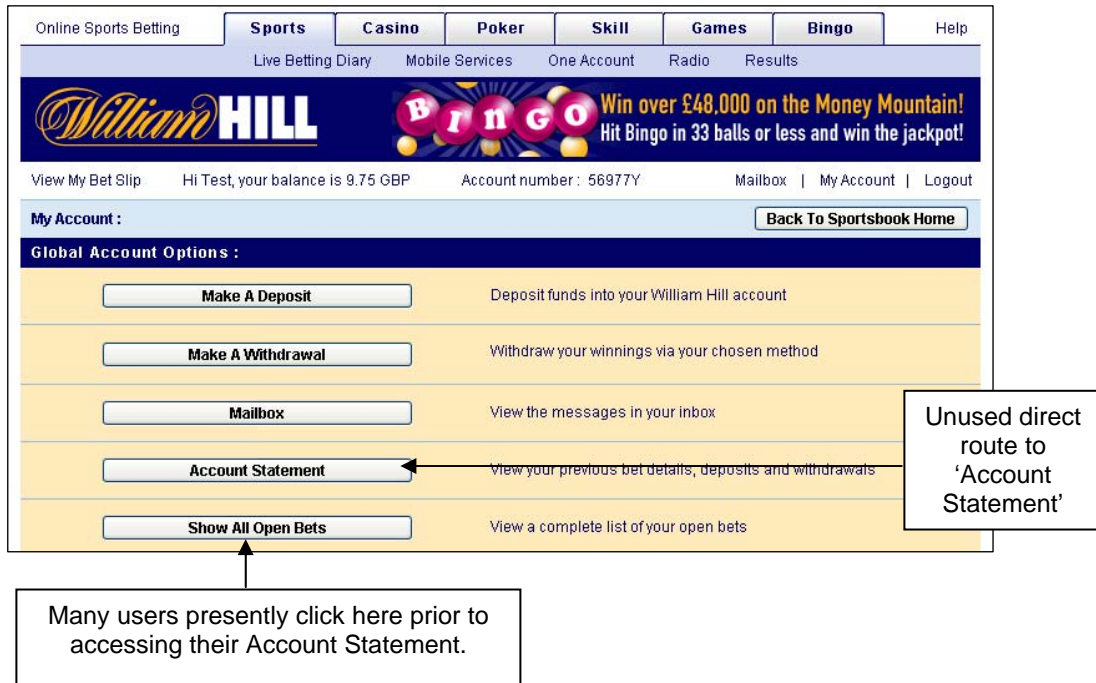
There are potentially many reasons for users accessing the Account Statement via this route. Users may be confused by the labelling of the buttons. Users may mistakenly click on 'Show All Open Bets' even when intending to choose 'Account Statement' because of the close vicinity of both buttons. Most likely is that users choose to view all their Open Bets, and then view their Account Statement to decide whether they can afford to place any more bets.

As shown in this example, the amount of information that can be inferred about user intentions is limited. This may be due to a limitation of the RedEye interface which does not include individual breakdown of time spent on each page within the pathway – which may help to indicate if links were intentionally clicked on or if users immediately navigated away. However, it is also a known facet of web analytics as discussed in published literature by Weischedel & Huizingh, (2006) and Hong & Landay, (2001) that it is impossible to infer users' exact intentions from pure data. Despite this limitation, raw data from web analytics is valuable in representing the whole user population. This type of data which shows the popularity and usage of pages is impossible to create in an observed environment, such as think aloud. Therefore this is where web analytics is superior to think aloud in generating representative data of realistic site use.

Having analysed the data from web analytics logs, these findings provide insight into areas of the website which are problematic and would be beneficial to explore further in the next stages of this study.

7.1.4 Findings to be used in further tests

The many error messages which faced users during registration were a concern, and resulted in a number of users giving up on registration to the William Hill website. The following controlled web analytics test, and think aloud evaluation therefore explored the process of registration in more detail with the hope of establishing the reason for these problems.

Although useful findings were gained through using the Red Eye interface to explore web analytics data, the experience was not flawless and will be critically discussed in the next section.

7.1.5 Reflection on the RedEye Interface

RedEye technology provides a great opportunity to gather data which represents usage of a website. Additionally, the criterion which may be selected from the RedEye interface allows the person creating the report to narrow down data to explore certain pages and pathways of the site which may be of interest.

There are however limitations of the interface which reduces the value of web analytics data when attempting to understand user experience.

It was difficult to make a start on analysis of the live data due to the quantity of data which covered the whole website. In this study I had input from Andrew Stockwell, who informed me of the client's interests in the site which was helpful to direct me in the analysis. Conversation with clients who may be aware of problems in the site, or areas which are critical to success of the site may be useful to highlight areas of the data which should be explored in more detail. If it is not possible to discuss details with clients, findings from exploration of the site by experts, or using very quick user tests may be used to uncover major issues with the site. These findings may then be explored in more detail within the web analytics to present realistic user behaviour within the pages of the site which are expected to cause difficulty.

The RedEye interface does not provide a breakdown of the number of seconds spent on each page by users, even though the full pathways are timed. The separate timings for each page would be useful to provide more information of what users may be doing on each page, and

how long they spend on each page. This may help analysts infer whether users are spending excessive periods on one page, and therefore may be having difficulties locating required information. Additionally, timings may help to decide whether pages are visited simply as a through route to another page, or even mistakenly if they are shown to only be viewed for a few seconds. Earlier in this paper it was mentioned that Shneiderman (1998) proposed web analytics as a method which may be used to examine the efficiency of regularly used aspects of a website. Although the total time taken to complete pathways may be used to infer this type of information, a more detailed breakdown of the step by step timings would be more useful. Despite this suggestion, it must be remembered that because web analytics data is collected without observation, it is impossible to know whether the timings presented by the data are true to the actual time spent by users on each page of a website. Users may be distracted whilst on pages, therefore time number of seconds logged as spent on each page may be much longer than the actual duration of time spent concentrating on each page.

Further restrictions of web analytics and use of the Red Eye interface are caused by limited tagging of the website. As mentioned by Weischedel & Huizingh (2006), web analytics data may not cover the whole site traffic because of untagged areas. At Red Eye, site owners can choose to tag either full pages, or tag individual links within pages. The latter option gathers more detailed data about not only the pathways taken by users, but also the specific links that users chose to access each page. This is more informative for analysts and can provide greater detail of the visibility of certain aspects of the website and the popularity of navigation links used. This information may also be useful to explore the success of promotional panels which may be used for marketing purposes. William Hill is tagged per page, rather than per link which limits the detail of information which may be collected.

According to Arroyo et al. (2006), server logs which present data similar to that shown in the RedEye interface, do not provide vital detailed information about what users are doing at a more granular level than the page-view level. Atterer et al. (2006) concur that information collected in classic log files needs to be more detailed and fine-grained to obtain a meaningful understanding of how users interact with web applications. The problem with such data is that information about how users actually interact with a web page and what they see is unavailable from such logs (Arroyo et al. 2006 and Atterer et al. (2006). As an alternative to page and link tags, eye tracking may be used to provide more detail of user experience when moving through a website. Modern eye tracking equipment now makes it possible to record and analyse; which elements are actually seen, where users look first and what users look at the most (Johansen and Hansen, 2006). Arroyo et al. (2006) also suggest that such data can be used

to quantify which sections of a webpage are read, glanced at, or skipped/ignored which can be vital in web evaluation.

Despite these advantages of eye tracking, data analysis can be very time consuming and the equipment is expensive (Arroyo et al., 2006 and Johansen & Hansen, 2006). Additionally, eye tracking equipment may reduce test validity – for example by notably slowing the system response or by requiring users to re-calibrate between tasks (Johansen & Hansen, 2006). Arroyo et al. (2006) also make the very valid point that eye tracking does not represent frequent visitors to a site, nor replicate their surfing conditions because the participant is aware of the testing and such tracking usually takes place in lab conditions. The fact that eye tracking can not be used in users' natural environment makes it impossible to use with web analytics. However, where the benefits of eye tracking may be utilized is during think aloud sessions to help determine whether users' comments truly represent what they see.

Having reviewed the use of live web analytics data in understanding user experience, the next phase of this study involved gathering and analysing controlled web analytics data.

7.2 Controlled web analytics data

The controlled web analytics study involved analysing data from users who complete tasks which were detailed in instructions emailed to them (see appendix 2). Broadly the tasks included registering with the site and then placing a bet on England to win the Ashes cricket tournament in 2009 and separately betting on USA to win the 37th Golf Ryder Cup. For both bets, participants were not required to enter a nominal stake onto the bet, but were simply asked to add the selections to the Betslip. The tasks were completed in participants' own environment and without observation. Data from eleven participants was analysed,

7.2.1 Registration process

The first instruction given to participants was to register with the William Hill website. The registration process required users to complete a number of fields relating to their personal details, however participants were offered fictitious details if they preferred not to use their own personal information.

Most pathways showed that users stepped through the full registration process with little difficultly. However, it was unclear from this data whether users did actually have problems during this process yet they managed to overcome them through deliberating over certain fields before realising what they were required to do to proceed. Additionally, this data did not

provide any information to site owners of how users felt about entering such personal data and whether they had any issues with the length of the registration process. This type of information could be derived from looking at the number of users who 'dropped of' mid-way through the registration process, however from pure data the reason for their leaving the site could be one of many.

The data did however show some problems which users faced during the registration process which are detailed below.

7.2.1.1 *Error message in registration*

Whilst entering personal details, three out of eleven users faced the same error message during the registration process. The data showed a code joined to this error, which Mr Stockwell informed me related to the Maximum Daily Deposit Limit field. This information was available when analysing the controlled web analytics data because web logs were provided by Mr Stockwell, directly pulled from data collected by RedEye technology. The details of the participants I had recruited were specifically pulled from the logs based on the Account Number allocated to each user. However, when analysing the live data in the previous phase of this study, path reports were used which do not include such detail about error codes. Path reports were used because when analysing data from the whole user population there was no access to their Account Number. The web logs from the controlled web analytics phase, showed the error in the data tables as an error message with a specific code relating to that field, shown in figure 12.

Figure 12: Security details error during registration

nourl=www.willhill.com/EN/SB/SelectUKAddress/ConfirmUKAddress
nourl=www.willhill.com/EN/SB/ConfirmUKAddress/SecurityDetails
nourl=www.willhill.com/EN/SB/SecurityDetails/error&errCode=SR3013
nourl=www.willhill.com/EN/SB/SecurityDetails/AccountConfirmation

Knowledge of this problem that users are facing relating to the Maximum Daily Deposit Field is interesting as a fact, however it does not give any insight into why users are having difficulty with completing this field. Having reviewed this field on the William Hill website I noted a number of issues users may have experienced, as shown in Figure 13.

Figure 13: Maximum Daily Deposit Limit field

Users may not realise this is a compulsory field

Users may not know how to activate the text entry box

* Maximum Daily Deposit Limit | Choose limit ▼ Daily Deposit Help

Choose limit
Default
£ 50
£ 100
£ 500
£ 1,000
£ 5,000
£ 10,000
£ 20,000
£ 50,000
£ 100,000
No Limit

*☐ I accept the Terms & Conditions , Privacy ... ite.

☑ Always credit free casino bonuses to my ac
☑ Keep me informed of future offers, events a

Back

Create My Account

English | Español | Deutsch | Italiano | Français | 中文

Users may not understand the terms 'Maximum Daily Deposit Limit' and may find the large figures within the option box daunting – therefore decide to leave this box empty.

There is a help link next to this field, which no user chose to access – even when having difficulty with this field.

The problems highlighted with this field may be representative of a limitation in the sample of users who took part in the controlled analytics study, who were mainly inexperienced in gambling online. As can be seen in Table 1, although five users occasionally used offline betting methods – only three users occasionally used online betting methods where the concept of the Maximum Daily Deposit Limit is relevant.

Despite some users facing this error message, all participants in the controlled web analytics test quickly overcame this problem and continued to successfully register with the site. This implies that this is not a critical error for task completion, but to improve the flow of registration – without the interruption of errors, consideration should be made to this field within the registration process. Conversely, during the live analytics analysis this was a frequent error which often resulted in users not completing the registration process therefore this issue may be more critical. This variation in results between the live and controlled web analytics sample

proves the benefit of the vast quantity of data logs which are gathered by web analytics, and may not be represented by only few users as used in the controlled exercise.

7.2.1.2 *Repeated registration process*

The instructions given to participants clearly specified that users should only register with the website once. However, four users out of eleven attempted to complete the registration process more than once. Although this fact was made clear in the data, it is difficult to infer why users did this.

One user logged out immediately after viewing the Account Confirmation page and then began the registration process again. This may have been done mistakenly, and then the participant may have believed that they hadn't successfully registered and then attempted to register again. If this suggestion is correct, it implies that the Account Confirmation screen is not clear enough and may allow users to logout too easily. However, this is just one proposition and it is difficult to understand why the participant may have done this without observation of actions and possible discussion with the user.

Two further users had reached the stage of registration where address details were required and then returned to the first page of the registration process and began the process again. This is an unusual path, and requires many assumptions to be made to come up with reasons for the actions. Again, without discussion with users it is difficult to infer the true reasoning of such pathways, and it is unwise to draw conclusions about actions without discussion with users.

Following registration, users were instructed to place a bet on specific Cricket and Golfing events.

7.2.2 Locating specific sporting sections and events

Instructions given to participants specified that initially users must choose England to win the Ashes 2009 Cricket tournament. Following this, participants were asked to bet upon USA to win the 37th Golf Ryder Cup. To simplify the tasks, participants were only asked to select these teams to win the events specified by adding these bets to their Betslip, without entering a monetary stake for each bet.

All users successfully located the Golf and Cricket sections of William Hill immediately after receiving confirmation of registering with the site. The results show that a number of users

navigated through to these sections by returning to the Home Page, often via the Make a Deposit page. Although this is not the direct route to choosing bet categories, this is the most realistic route for users who, if were not participating in a controlled study, would expect to make a deposit into their account before betting.

A design recommendation which may be made purely from this data is that the left hand navigational bar, which is present on the home page may be included on this Account Confirmation page to enable users to proceed with betting more easily, as shown in the mocked up image in figure 14.

Figure 14: Design recommendation to include left hand navigation side bar
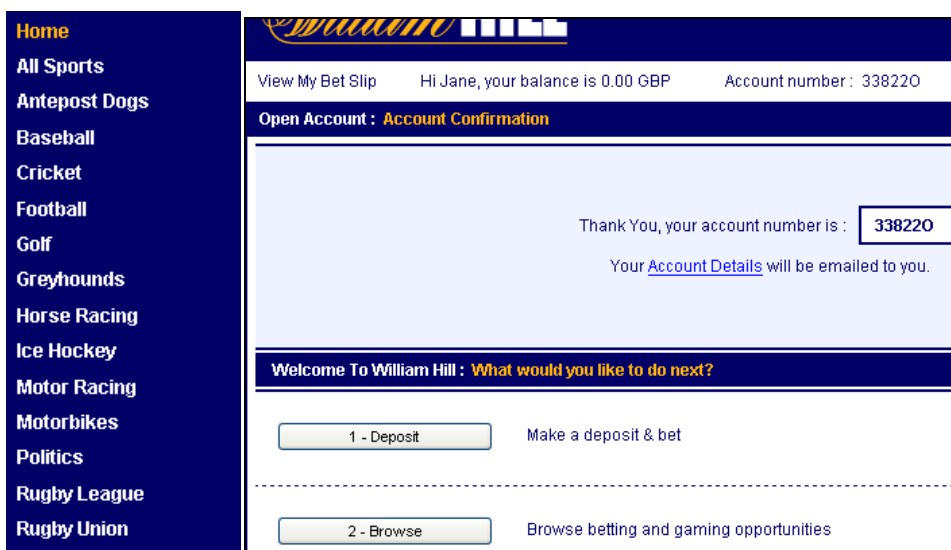
7.2.2.1 *Choice of correct event*

Instructions given to participants specified which team they should have bet would win both the 2009 Ashes tournament and the 37th Golf Ryder cup. However, it is unclear from the web analytics whether users correctly chose the specific event and the team users bet to win. This is a limitation of the RedEye analytics logs, in that the data shows the pages visited by users but not the exact data entered into each page. It is clear that all users accessed the Golf and Cricket sections of William Hill, as requested to in the instructions – however, the data does not record further details of the event and the nature of the bet placed. This is a major limitation of the web analytics collected for William Hill, which enables users to place various types of bets on many different events – all within the same sporting category. In a meeting with Mr John Dumas of optimum.web, it was made clear that each company who use web analytics data

choose the level of detail they wish to obtain from the analytics reports. John Dumas highlighted that the technical ability to collect such detailed information is available in the present web analytics technology, however it is expensive and time consuming to apply therefore many companies, including William Hill, choose to implement only limited capabilities. Although this is a limitation of the RedEye interface, due to the vast range of web analytics services available by numerous suppliers it is unclear whether this is an inherent limitation of web analytics as a whole or only of specific companies' packages.

7.2.3 Online Help required

Four users accessed the Help feature of the website when attempting to place a bet. All that is represented in the data is users accessing a help screen, which does not provide an indication of the type of help they required. What is shown however is that two users chose to view a help screen within the Cricket section and two other users accessed help within the Golf section. This suggests that the help required by users was not sport specific and related to the general use of the website to place bets.

Having used the Help function, all four users went on to successfully complete the tasks. This suggests that the help feature within William Hill was successful in aiding users to place bets on the site. However, what can not be inferred from this data is why users required help, and therefore how the site may be improved to ensure users can move through the site without requiring additional assistance. This type of information may only be captured through web analytics if users were to search a help database of various help categories, rather than only being presented with general help about the given page.

Having reported all the findings gained from the controlled web analytics study, the following section reflects on this phase.

7.2.4 Reflection on the controlled web analytics study

The controlled analytics results provided me with data which was useful to show where users had obvious problems e.g. incurring an error message and accessing the help function. However, this data was also represented in the live analytics data therefore the controlled analytics phase had limited value in this study.

Knowledge of user objectives during the controlled analytics phase did not provide any greater insight into user experience of participants. During the initial task of registering with the site, it was unnecessary to know user objectives. This is because when the page shown in the web

logs as 'StartedRegistration' is present it is fairly clear for analysts to infer from the data that users intend to register with the site. Data referring to registration was therefore equally useful from the live data analysis when objectives were unknown to where objectives were known in this controlled part of the study.

However, to determine whether users had successfully placed bets on the events specified in the instructions it was necessary to know details of the tasks set. The web analytics data did not provide a sufficiently detailed breakdown of the selections made within the placing bets stage of the test therefore knowledge of such user objectives was unhelpful. As mentioned previously, this may be because William Hill only use page tags, therefore data referring to the individual links chosen within pages was not recorded. In determining the exact details of users' bets, this information would be required, however it is perhaps not necessary for collection to meet William Hill's objectives from the web analytics data. If William Hill are mainly interested in whether people get stuck on particular pages, or where users leave the site then link tags are not required, and page tags will suffice.

The study of both live data and controlled analytics data has proven that there is potential to infer some user objectives simply through viewing pure live data, without knowing users intentions – as was done in the controlled exercise. In fact, using the live data was even more beneficial than the controlled data because it provided data of pathways representing the whole user population, rather than only a small sample of eleven users. Despite this, as found in the live data analysis it was difficult to know where to begin analysing the data because of the quantity and the vast spectrum of pathways it covers. Due to this, it was useful to have objectives as to what areas of the site are of interest before beginning analysis.

Following complete analysis of the web analytics data, the think aloud phase of the study began.

7.3 Think aloud

The think aloud sessions involved users completing tasks on the William Hill site whilst verbalising their thoughts and being observed by myself. The following section details the findings from the sessions which took place with five participants.

7.3.1 General comments

Overall feedback of users' experience on the William Hill website was mainly positive. Users had few problems navigating the site and many commented on the great array of events

covering various sporting categories which may be bet upon. During registration there were a few issues which negatively affected the user experience, and might have resulted in users leaving the site prior to completing registration if they had been in a more realistic situation.

The following sections detail the findings from the think aloud sessions.

7.3.1.1　Home page

Figure 15: Existing home page of William Hill website
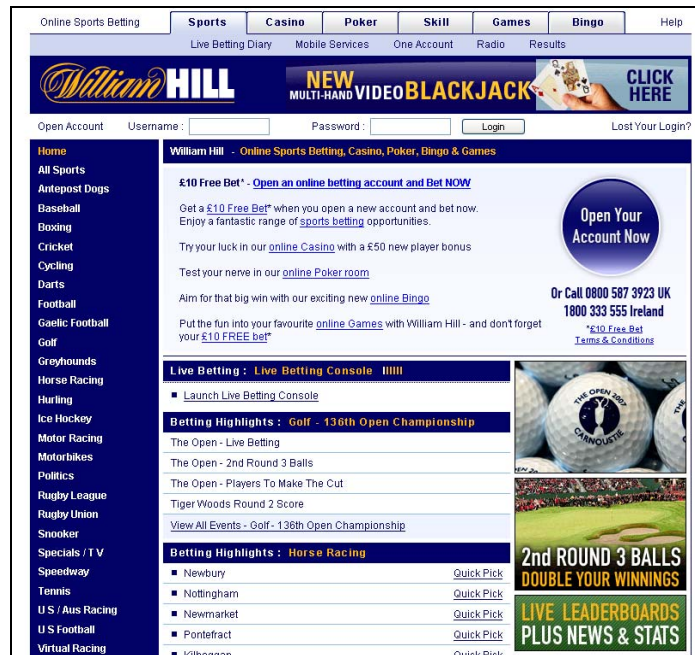
The home page shown in figure 15, generated mixed feedback from all users. Although some users commented on the business of the home page, they felt that this was seemingly worse because they were not familiar with betting. Additionally, although they commented on the vast number of links down the left hand side, this was seen to be positive by all users because it showed the wide catalogue of events gamblers could choose from and the listing used clear categorisation by sport type.

"This page looks ok, if I was a gambler I might know where to go a bit more…but I recognise all these sports down the left hand side, so would choose one of them"
(User 1)

"I don't like it…it's too confusing…it's too much….too complicated….all the different boxes, I think that's why I chose Ladbrokes before because I couldn't find what I wanted" **(User 3)**

"It's really busy this page – but I guess they've got lots of things you can bet on…you can see all the sports down the left hand side which is good – but it's all a bit blue and white!" **(User 5)**

7.3.1.2   Locating login buttons

From the home page, the final task of the think aloud session involved users attempting to log back into the William Hill site, using the registration details they had created earlier. Many users went about this differently, which highlighted issues with this section of the home page. Figure 16 shows the Log-in section of the home page.

Figure 16: Log in section of the home page of the William Hill website.

The clarity and differentiation of the four sections; Open Account, Username, Password and Login button within this upper bar of the home page is not clear. Two users assumed that the Login button would take them to a page with text entry fields to enter their username and password – completely overlooking the Username and Password boxes on the home page. This assumption by participants is in line with the findings from the live web analytics which highlighted many users who faced an error at this stage of login. Furthermore a number of users mentioned that these boxes were not clear on the home page.

"I didn't really notice those boxes, I just spotted the button that says login and went there….it seems like all those things are the same [Open Account, Username, Password]…they're all the same text….but they're completely different….some should be bold or something to make them look a bit different". **(User 1)**

"Those boxes are clear,, but I think the whole page is too busy and has advertisements all over it – they should be in a box on the left or the right and separate from everything else". **(User 3)**

"I wouldn't include the username and password boxes on this screen...there should just have the sign in button as that's what I saw first – or maybe just have an enter next to the boxes"
(User 4)

"Everything blends in a bit too much on this page…I saw the Login button, but now I've seen the boxes it is quicker just to put your username and password in here…but those boxes should be made more clear" **(User 5)**

7.3.2 Registration

The first task given to users was to register with the William Hill website. Users were able to choose whether they used their own personal details to complete registration or used fictitious details which were still realistic.

7.3.2.1 Username error

Part of the registration process required users to create a unique username. Four out of five users created a username which was already taken by another William Hill customer, therefore they faced an error message which offered other usernames which were similar to their original choice. Although it may be seen as helpful to offer similar usernames, it is likely that these would be forgotten by users when they returned to the site to log back in. Furthermore the error message was infuriating for users, and frustrated them when trying to complete the registration process. Three users who faced this problem suggested that using their email address as a username would be more suitable because it is definitely unique to each user and is unlikely to be forgotten.

"I think the username should just be your email address then you would never have this error and you would never forget it" **(User 3)**

"What I hate about these websites is when you have to think of a username…why don't they just use an email, then you don't have to think of something every time which I always forget"
(User 4)

"I always forget these usernames you have to make up…especially when mine is usually already taken….it should be something like your email address, then you would never forget"
(User 5)

7.3.2.2 Password error

All users came across an error message during the registration process. Some error messages related to usernames which were not unique, others regarding the Maximum Daily Deposit Limit. However, regardless of the type of error which appeared on this page, users were also faced with an error message relating to their password – which had in fact met all criteria.

"Ok, so my username already exists…that doesn't surprise me…but it looks like I need a new password too – so I'll type a new one" **(User 1)**

"I just did a password which was six characters, why have I got to do it again…oh, I feel annoyed [at error message] – I've just been through all of this and now they don't like it and I have to enter it all again" **(User 2)**

"Why do I need a new password….I've already entered it" **(User 2)**

"My username already exists….I don't know why I've got an error about the password – t-e-s-t-1-2-3 that was 6 characters" .**(User 3)**

There was no reason for this error message to be appearing and it resulted in frustration for users. Not only did this annoy participants but also users may have entered a new password which they were not familiar with, therefore would have been more likely to forget it when logging in, in future.

7.3.2.3 Maximum Daily Deposit Limit

Maximum Daily Deposit Limit is a field which users must fill in as part of registration. Only one user immediately understood what was meant by the term Maximum Daily Deposit Limit. Four out of five participants used the help feature, which they all found helpful in explaining the term. Figure 17 shows details of this field, as well as the help screen which users accessed.

Figure 17: Maximum Daily Deposit field and help feature.
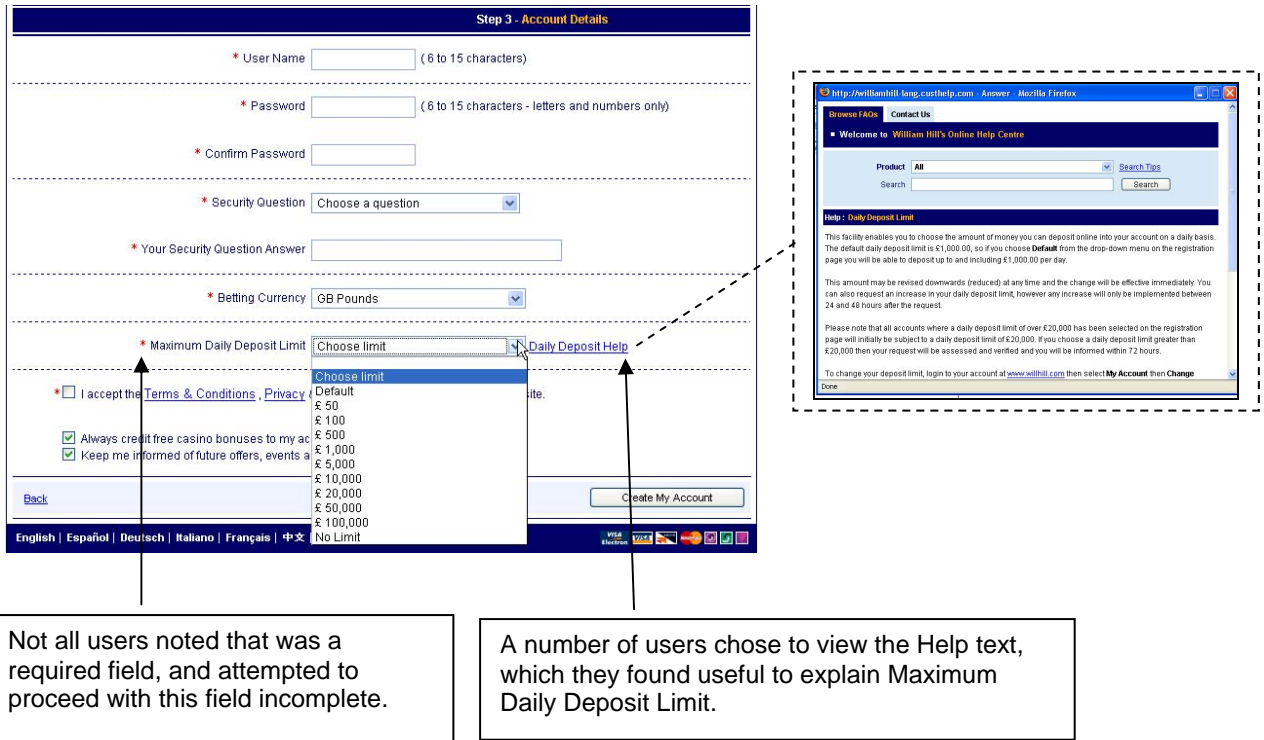
Not all users noted that was a required field, and attempted to proceed with this field incomplete.

A number of users chose to view the Help text, which they found useful to explain Maximum Daily Deposit Limit.

"I think its how much you're going to put into your account on a daily basis…so I can't spend more than a certain amount in one day?" **(User 2)**

[After viewing Help] *"It's quite a good idea actually, to stop people spending all their money on gambling"* **(User 1)**

"I think it's the amount I can put in daily to this account…I would choose the smallest amount as I don't want to spend much…but can I change it if I want to in the future?....I would check the Help thing to find out" **(User 3)**

"You can select how much you can bet each day which is probably quite a good idea" **(User 4)**

"I guess it's something about how you deposit money into your account…but I don't really know…there's Help though, so I'll look in that" **(User 5)**

Although most users understood what was meant by the term after looking in the Help feature, two users did not feel that this was relevant to them – because they were not frequent betters

and therefore did not need to set a limit. This resulted in users leaving this field incomplete having not noted it being a 'Required field' and therefore facing another error message. Additionally, it was mentioned that the large figures listed as Deposit Limits may alarm less frequent and thriftier users of the site.

"This is assuming you bet everyday – but I'm just an occasional better…this would frighten people I would have thought…especially with such large amounts of money they're talking about" **(User 2)**

"I don't want a maximum daily limit, I'm an occasional better and I don't' want to have to enter one….can I just write 'No Limit'?" **(User 2)**

"But I know I'm not going to spend too much money on this website so I don't need to fill it in" **(User 5)**

When users voiced issues regarding this field of registration, they also made recommendations of other terms which may better describe the Maximum Daily Deposit Limit concept without requiring users to access Help to aid their understanding.   Users came up with the following recommendations:

"Maximum Daily Expenditure" **(User 1)**

"Maximum Daily Betting Expenditure" **(User 2)**

"Maximum to be spent on betting per day" **(User 5)**

7.3.2.4   Security question

Requiring users to enter a security question during registration did not make them value the account as any more secure. Three users complained that these types of questions were used too frequently and occasionally to access bank accounts which made the consequences of someone finding out the answer more dangerous. It was also commented by two users that the answers to questions listed in the drop down box are fairly simple for people to find out and therefore do not add to the security of the account. Figure 18 shows the Security Question field which is a mandatory field for users to complete during registration.

Figure 18: Security Question field.

"I don't think it makes it any more secure….if someone wanted to do identity fraud or something they could easily find out my Mum's maiden name…" **(User 1)**

"I don't really want to tell anyone this information....these are all the details I use for my bank accounts so I try to tell few people…the more people who know it the more likely people are going to find out". **(User 2)**

"….I don't like having to add this kind of thing [security question] *as that's what they say on bank accounts…I would try to choose the most obscure one so no one knows my bank details".* *(User 3)*

"They're kind of standard [Security questions]*…it's good actually cause I'm rubbish at passwords"* **(User 4)**

"I'm not sure why they need to know my Mum's maiden name…everyone seems to ask for this these days – but I don't think it would be that difficult for a fraudster to find out" **(User 5)**

7.3.3 Placing bets

As part of the think aloud session, users were asked to bet on England to with the 2009 Ashes tournament and USA to win the 37th Golf Ryder Cup. Users were asked to do this immediately following registration.

7.3.3.1   Locating sporting events

No user had difficulty finding the specific sporting event they were asked to place a bet on. Furthermore, a number of users commented positively on the left hand navigation list of the various categories which may be bet upon.

"It's very easy to navigate, you know which sport you want to bet on and choose it from the long list on the left and then it has a wide selection of events which you can choose" **(User 1)**

"They're clear on the left hand side, you just click on them and you're in". **(User 3)**

"That was easy to find cricket….and it looks like you can bet on loads of different things from this list [left hand navigation]*"* **(User 5)**

7.3.3.2   Selecting events

Having located an event to place a bet on, users were unsure how to proceed. Most users expected to be shown further details of the event and options for bets on that event when selecting the tick box next to the event name, as shown in figure 19. Users were unsure what Show Selected Markets meant, and often did not notice this button until they had realised that they had to take further action after ticking the box.

Figure 19: Selecting event page on William Hill website.
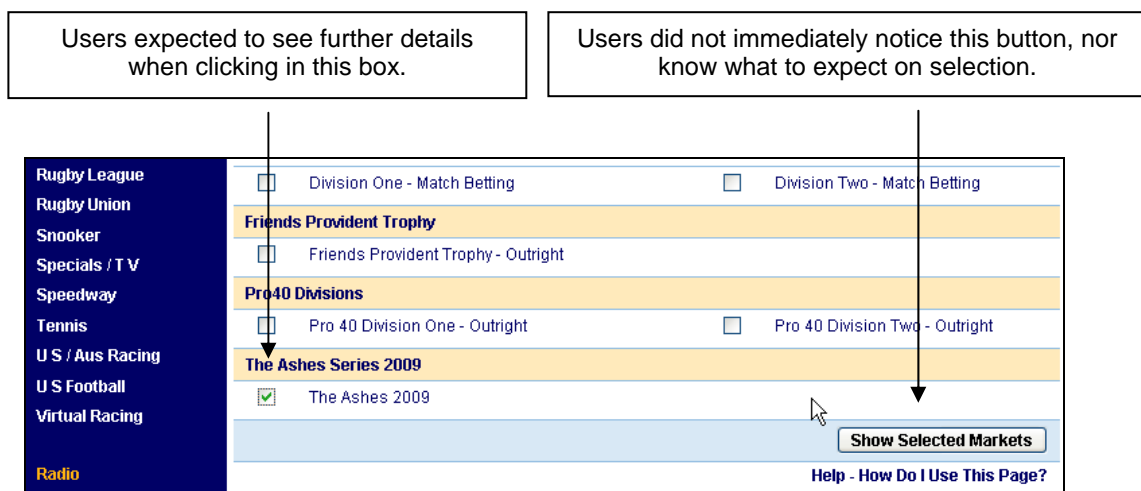
"I'm not sure what Show All Current Markets means – but there's nothing else I can click on"
(User 1)

"I thought something would come up when I ticked that box….oh I can see Show Selected Markets but I've no idea what that means" **(User 2)**

"I'd tick that [waits for response]*…no, I don't know what that tick box was for….actually there's a hyperlink, I would rather click that [clicks hyperlink]"* **(User 3)**

"[User ticks box]…So I've chosen the Ashes 2009…what do I do next….how can I say I want England to win?" **(User 5)**

7.3.3.3 Home Vs Away

When using instructions to place a bet on a certain country (e.g. England) to win an event, some users did not understand the Home and Away labels which were used to represent the teams involved in the event, shown in figure 20. Three of the users who had difficulty realised that the Home team represented the team playing on their 'Home ground', however they did not know the location of the event and therefore did not know which was the Home team.

Figure 20: Bet score selection page

Bet Finder :	Please Select a Coupon	▼		Decimal Odds	▼
Cricket : The Ashes 2009					
Bet Until : 11:00 25/12/2007					
Match		**Home**	**Draw**	**Away**	
England v Australia		3.25 ☐	5.50 ☐	1.57 ☐	

"Home is 3.25, but I don't know where it's being held…maybe Home is England because I'm in England now?....I know home is where the event is being held, maybe if I knew more about cricket I would know where the Ashes 2009 is on?" **(User 2)**

"I'd much prefer it to be clear and just say England or Australia to win" **(User 2)**

"Match, Home, Draw or Away…I have no idea, I don't know what to do! I suppose it's played in this country isn't it?....I don't know where it's played, I haven't got a clue!…why doesn't it just say England or Australia to win, I wouldn't bother with that, I would just close it and try somewhere else!....I might click on betting Help but to be honest I would just close the window"
(User 3)

"I know because England is written first they're at home…but I don't think some people who don't know about sport would know that" **(User 4)**

"So, I'm betting on England to win?....it says Home and Away – but where is the Ashes playing in 2009…is there somewhere I can find out where it's playing so I know whether England is

Home or Away?.....that's really confusing though – since everyone knows which country they
want to bet on, why don't they just have an England box? **(User 5)**

This finding may be representative of a limitation of the sample used in the think aloud study, who are not regular gamblers. However, even one of the users who does bet online had difficulty with this aspect of betting. Additionally, with the influx of gambling opportunities in more widespread fields; such as Big Brother the terminology on the site should be simplified to ensure it is understood by the widest user population possible.

7.3.3.4 Decimal and fractional odds

A number of users commented on the use of decimal odds when viewing details of the event to bet upon, as shown in figure 20. All of these users guessed correctly the meaning of the decimal odds, however when they changed the format to fractional odds all users found this more clear and the expected format of odds.

"It's 3 point 25 to home win…I presume it's £3.25….maybe online betting stakes are a bit more
complicated than traditional ones….[notices fraction/decimal option]…..look, you can change
the odds to fractions,….that's much more traditional" **(User 1)**

"So, for every £ I put on I would get £3.25 back…I understand odds, but that's different to what
I've seen before….[changes to fractional]…that's better, that's what I'm used to from the dogs"
(User 3)

"3.25…what's that?...normally it's two-to-one or something…I'm not really sure what 3.25
means" **(User 5)**

From this user feedback, it is recommended that William Hill present fractional odds as a default, which users may alter to decimal format if they prefer. Having found that most users expect to see odds as fractions, it would be best for this to be format used when users first see the odds – rather than requiring them to change the format of the odds before they can easily understand them.

Having reported all the findings gained from the think aloud sessions, the following section reflects on this part of the study and its value in understanding user experience.

7.4 Reflection on Think aloud process

In line with literature published by Ramey et al. (2006) the think aloud sessions generated a wealth of results and provided the greatest insight into user experience whilst using the website. This covered general comments about the site's functions and look and feel, along with detailed issues about specific parts of the website.

All users conformed to the think aloud protocol well, making it less necessary to prompt users to verbalise their thoughts. Additionally, as noted by Preece et al. (2002) users' affective movements were indicative of usability issues as they hovered or paused before taking action. Although these actions did not specifically provide details of what users were having difficulty with, this was where users required prompting by the moderator to describe what they were thinking. It was important to see these kinds of actions and prompt users for comments, because without such visual cues users may not verbalise their concerns and may eventually work out how to continue with using the site – without thinking aloud and therefore without the moderator taking note of the issue.

There was no evidence to suggest that the increased cognitive load of users caused by thinking aloud whilst working on the website affected the success of task completion, as suggested in much literature by Preece (1994), Ramey et al (2002), Preece et al (2002), Neilsen (1993). This may have been due to the fairly simple nature of the tasks used in the think aloud sessions. Where the effect of increased cognitive load may devalue the accuracy of results is when timing users to complete tasks – because the time taken for users to work on the website whilst verbalising their thoughts is likely to be greater than the realistic timings.

Since all data gathering and analysis was completed by myself, inconsistency issues discussed by Ramey et al (2006) and Hertzum & Jacobsen (2001) were not explored in this project. This is a limitation of the study, however because all sessions were run by the same moderator, the findings from the test should have been gathered in a consistent manner. Additionally, the results were analysed by myself which ensures that the analysis was consistent. However, although I attempted to maintain a neutral position throughout the sessions, it is likely that I influenced the sessions and analysis to focus on issues within the site that I found particularly interesting.

Furthermore, having analysed the web analytics data and come across issues with the site which didn't have clear reasoning prior to the think aloud phase, during the think aloud sessions it was tempting to explore these areas in more detail to infer what users may have

been thinking at these points. This is inline with Norgaard and Hornbaek's (2006) concern of usability practitioners confirming known issues during think aloud sessions.

Despite this potential contamination of findings, the results of the think aloud studies were in line with the findings of the web analytics data, which suggests that the evaluation was accurate with minimal influence from myself. Where my influence may have affected results was when users noted a problem on the site I may have explored the areas which came up in the web analytics tests in more detail in an attempt to determine what users were thinking at those points. Although this may have reduced the purity of the results it has also provided me with experience in combining both methods to draw out more useful findings for the client, which may affect more than only the five participants of the think aloud.

Having completed and analysed all three phases of this study the results were then combined to explore where each method may be best used independently and whether they may be combined to better understand user experience.

8 Discussion of results

The following section details the unique findings from both usability evaluation methods explored in this paper: web analytics and think aloud. This is useful to establish the differences between the types of findings gained from both methods, and whether either method is more useful in determining certain usability issues. This also helps to produce guidelines determining when each method may be best used throughout the User Centred Design process, and whether both methods may be used to complement each other during web evaluation to gain a deeper insight into user experience.

8.1 Unique findings from web analytics

As already discussed in detail in the results section of this paper, web analytics generated much quantitative data which represented the whole user population of the William Hill website.

Such quantitative data is valuable in web evaluation because it provides detail of all user pathways taken throughout the website. This is particularly useful for website owners when reviewing the navigation and information architecture of sites which may be altered based on trends in user pathways.

Due to the fact that web analytics is collected when users are in their own environment and surfing naturally, the data is useful to explore popularity of pages. This could be particularly useful for e-commerce sites to determine which products are viewed most regularly and to spot any trends in user behaviour when exploring items for sale.

In this study, web analytics data was particularly useful to highlight error messages faced by users of the website. This is important for website owners to be aware of so that they have some indication of where users are having difficulty and therefore where changes may be required to improve user experience. Knowledge of error messages may also be useful if website owners are unsure why users fail to complete transactions – providing them with clues of the difficulties users face.

Finally, data produced by web analytics is useful to highlight drop off points of users. This provides website owners with clear information of how many users are leaving the site at various places within the site. There could be many reasons for users leaving a site, however trends in this data may indicate problematic areas of the site which may be rectified to extend users time spent on the site. This information of drop off points is impossible to accurately gain from think aloud studies because of the controlled and observed environment in which participants use the site.

The unique findings discussed can be divulged accurately only through web analytics and not think aloud because web analytics data represents the whole user population, and is collected unobtrusively.

Having explored the unique findings gained through analysis of the web analytics data, the following section details the specific types of findings gained through think aloud evaluation.

8.2 Unique findings from think aloud

The think aloud studies provided great insight into users' attitudes through general comments made by participants throughout the sessions. Such topics include comments about the look and feel of the site, or issues with the site which do not result in a usability catastrophe, but which may be useful to be addressed in future rounds of testing. Additionally, during the think aloud sessions a number of users expressed security concerns, for example when answering the security question. This type of information is vital for many transactional websites such as William Hill, where customers may spend a vast amount of money online and therefore trust is

paramount. Such detailed information about security concerns is of less value when evaluating information based websites. However, general comments made by participants about comprehension issues during think aloud sessions may be critical to the success of information based websites.

In addition to general comments made by participants, think aloud evaluation also provided detailed insight into thought processes' of users as they used the website. This included information about users' expectations of certain functions of the website along with perceptions of specific aspects of the site. In this study, problems which were uncovered around terminology used on the William Hill site, for example on the Maximum Daily Deposit field, were important to gain an insight into users' expectations and understanding. As shown in this study, usability practitioners can combine users' comments with their actions to uncover usability issues. However, a drawback at this stage is that interpretation by the usability practitioner is required, which may vary between individuals and therefore reduce the purity of results as discussed by Norgaard and Hornbaek (2006) and Hertzum & Jacobsen (2001).

Having established the different types of findings which are gained through independent web analytics and think aloud studies, the next short section details corresponding results which were found using both methods.

8.3 Overlapped findings

Findings from both the web analytics and the think aloud studies uncovered areas of the William Hill website which were problematic for users and presented an obstacle in their journey on the site. Specific issues highlighted in all studies of this project included error messages faced by users during the registration process on the William Hill website. This overlap shows that both methods may be used independently to correctly derive similar usability issues. However, the results of the web analytics provided little detail of the problems faced by users and therefore give limited insight to usability practitioners of user experience. The results of the think aloud sessions however, provided specific detail of where users had problems through observation of participants' actions and their comments. Additionally, the think aloud sessions generated further results through users' conversational comments which indicated further usability issues which were not highlighted in the web analytics data. However, the data from the think aloud sessions was limited because it only represented five potential users of the site, who were not

necessarily similar to William Hill customers – and used the site whilst being observed in controlled conditions.

8.4 Application of Web analytics and Think aloud

Having explored the unique findings gained from web analytics and think aloud evaluation methods, it is useful to determine where each of these methods should be best used in any web design project. Although this paper has concentrated on the Evaluation phase of design projects, it is clear that such methods may be useful at further phases of any User Centred Design project.

User Centred Design (UCD) is a broad term to describe a design process in which end-users influence how a design takes shape (Abras et al. 2004). According to Preece et al. (2002) real users and their goals should be the driving force behind development of a product during a User Centred Design project. The specific phases at which users may be consulted during this type of project is not specific, however Figure 21 represents a widely adopted UCD process.
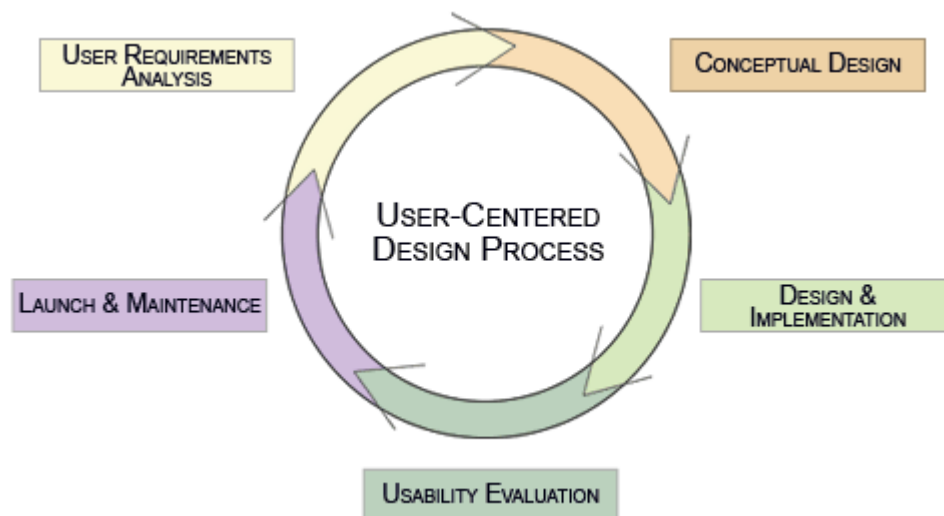
Figure 21: User Centered Design process.
Taken from the Usability and Accessibility Centre (2007).

During this process think aloud may be used at various stages including Conceptual design and further Design, to gather first hand user feedback on designs. Think aloud

sessions may be run using paper prototypes which represent early design ideas. The process of paper prototyping may be used as a communication medium through which users can articulate their requirements, which may have been overlooked at the initial stage of User Requirements gathering (Scanlon, 1998, Preece et al. 2002 and Rudd et al, 1996). Validating design ideas through user feedback also ensures real end user input in the design and iteration of the interface (Snyder, 2003), as advocated in the User Centered Design approach. The method of think aloud is most suitable to this phase because such early prototypes are not live and therefore inaccessible to real users of the site. This means that web analytics can not be used at this phase of the UCD process.

In addition to using think aloud in the early stages of a UCD project, as found in this study this method is also of great use during Usability Evaluation to uncover all usability issues which may affect user experience. This method may also be used intermittently during the Maintenance phase of a website project to find out if there are any usability problems which may affect the live user population, yet were not uncovered during the Evaluation phase.

In the User Centred Design process of a new project starting from scratch, there is little role for web analytics, except during the Maintenance phase as described above. Throughout this phase web analytics data can be used to monitor user pathways by the whole user population and determine where usability issues occur and how much of the user population is affected by such issues. This information may then be used to base scenarios for think aloud testing, to explore the issues uncovered by web analytics in more detail.

So far, the discussion has concentrated on new User Centred Design projects which are not based on existing websites. However, many UCD projects are actually redesigns, where an existing website is being revised. In this case, both think aloud and web analytics may be valuable in various stages of the User Centred Design cycle. Figure 22 shows an adaptation of the traditional UCD diagram displaying where both methods may be applied to truly understand user experience, based on findings from this study.
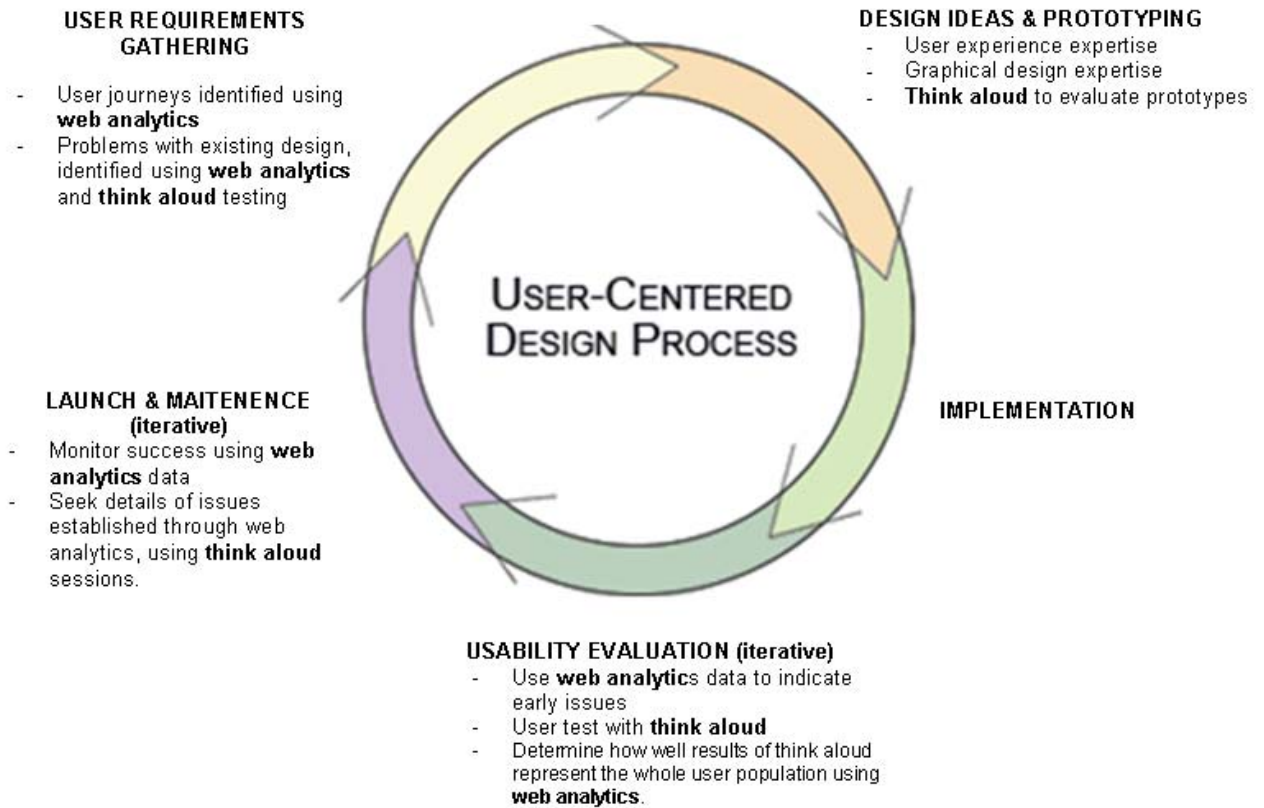
Figure 22: Application of web analytics and think aloud during the User Centered Design process. Adapted from the Usability and Accessibility Centre (2007).

As shown in figure 22, think aloud studies and web analytics data may be used in the User Requirements gathering phase of a redesign UCD project to understand what users need when visiting a website. Following this, think aloud sessions may be run on the existing site to uncover usability issues which must be addressed in the redesign. Initially, web analytics data may be used to determine which journeys are being taken by current users of the site. This information is very valuable to feed into the early stages of a redesign project to ensure that all current user needs are met. Further methods may be used to address the needs of users who do not currently use the website, including surveys and focus groups to determine why they may not presently use the site and which features may be added to the site to make it more appealing.

Web analytics data may also be used in the early User Requirements stages of a redesign UCD project to determine any critical issues with the existing website which hinder successful use by customers. Web analytics is particularly useful at this stage because of its representation of the whole user population. Additionally, this may be important in projects

with limited budgets ensuring that only issues which are proven to affect current users of the site are concentrated on in the redesign, rather than spending valuable time exploring potential usability issues which do not affect user experience. Additionally, web analytics data may also be used at the end of a UCD project to establish the success of any new designs and compare data from prior to the changes with data during Maintenance phases to ensure continual awareness of user experience.

Despite these ideas of how web analytics may be used early in any UCD project, as found in the web analytics part of this study, it is difficult to know where to begin drawing conclusions from such a vast range of data. If time and budgets are unlimited then more time may be spent analysing the full data logs without influence of client or expert opinion about potential usability issues at this early stage of analysis. However, where time is short, consultation with clients may be beneficial to understand any specific areas of the site which are critical to successful user experience and meeting strategic objectives of the client. If this type of information is unknown by clients, expert evaluation may be used to identify any potentially difficult areas for users. Following this, the potentially problematic areas uncovered may be explored in detail within the web analytics data to determine whether the real user population face the problems suggested by experts.

As suggested by Weischedel & Huizingh (2006) and Hong & Landay (2001), web analytics data is limited in providing information of why users took the pathways they chose and how they found the experience of using the site. This was noted also in this study, however more detailed information about user experience was uncovered during the proceeding think aloud sessions. This suggests that where web analytics data is limited and can not be used to draw accurate conclusions of user experience, think aloud sessions may be run to derive from participants the potential reasons for such problems. However, basing think aloud studies on findings from web analytics data may increase the concern highlighted by Norgaard and Hornbaek (2006) around usability practitioners who run think aloud sessions, using participants to confirm known issues. Although this may be seen negative, expert hypothesis and evidence from sources such as analytics must be used when planning and conducting testing, in order to focus on areas which need attention. In a meeting with John Dumas of optimum.web, he stated based on many years experience of running think aloud studies, that no test can be totally purist with an unlimited amount of time, due to budget and user fatigue. Mr Dumas did however also

point out that a balance must be made between usability practitioners' attempting to prove or disprove hypotheses, and gaining purist data from participants.

Despite Norgaard and Hornbaek's (2006) concerns, as found in this study web analytics data does not provide details of the reasons for usability issues and only presents pure data to usability practitioners. The reasoning information about why issues occur may only be divulged through observation and discussions with users during think aloud sessions. Additionally, knowledge of issues which are found in the web analytics data may help to ensure that think aloud sessions are run efficiently, exploring only errors which are facing a notable sample of the real user population.

This proposal to use the results from both web analytics and think aloud together is beneficial at any point of the UCD process because it uses a combination of data which represents the true user population as well as results from think aloud studies which produce much qualitative data – therefore gaining a deep understanding of user experience.

Following implementation of designs, the Usability Evaluation phase attempts to uncover all issues which compromise user experience. This paper concentrated on web analytics and think aloud methods, and the results of the studies helped to uncover how both methods may be used to better understand user experience; particularly during web Usability Evaluation.

Norgaard & Hornbaek (2006) and Hertzum & Jacobsen, (2001) highlight two main concerns regarding the think aloud method. Firstly uninstructed usability practitioners running think aloud sessions and secondly inconsistent analysis of findings, both of which reduce the purity of results from this method. To overcome such issues, this study proposes that web analytics data, which represents users in their natural environment and requires less interpretation by analysts, may be used in conjunction with the findings from think aloud tests to check whether the results of think aloud sessions are truly representative of user audience. However, as found in this study it is sometimes necessary for analysts to make assumptions about users' intentions to make sense of analytics data. It is therefore important to only use data which clearly shows user issues, with minimal requirement of analyst interpretation. This ensures that inconsistencies of interpretation both in the early analytics analysis and think aloud analysis are not combined.

To extend this study, future research may look into the suitability of combining eye tracking technology with data from web analytics and think aloud, as an additional method to deepen

understanding of user experience. As discussed previously, because eye tracking data is collected through specialist software which participants must have access to, it is impossible to use in combination with web analytics, which is presently collected unobtrusively. There is potential for this method to be more suitably used to validate results from think aloud sessions. Despite this suggestion, the limitations of using a controlled methodology where the participants are aware of observation and collection of eye tracking data would still present limitations – therefore it would be interesting for further research to explore the value of applying eye tracking technology during think aloud sessions.

Further issues around the 'evaluator effect', published by Hertzum & Jacobsen (2001) were not directly explored in this study because only one usability practitioner facilitated all sessions. However, the insightful findings provided by web analytics were in line with the findings from the think aloud sessions which suggests that the data gathered from both methods was accurate in representing user experience. In future research, the evaluator effect could be explored in more detail though the use of further usability practitioners who would run and analyse the sessions independently. This would uncover the level of any discrepancies between conclusions made by the various practitioners and therefore how necessary web analytics may be in validating results.

Using data from web analytics to validate findings from think aloud sessions would also reduce the controversial effect of limited numbers of participants in think aloud sessions, and how well these users realistically represent the whole user population, as discussed by Wixon (2003). There is however a mismatch of expertise in participants of think aloud studies and users represented in web analytics data. The live data shows the pathways taken by real and potentially experienced users of a website however think aloud testing may use novice participants, depending on the specification used to recruit users. Although there is a discrepancy between participants' experience it is important that when analysing the findings of think aloud sessions, the web analytics data which represents the real user population is taken into account to minimise participants' potential limitations in not being truly representative of the user population.

Although the suggestions made in this paper, advocate combined use of web analytics and think aloud to understand user experience, further research may be valuable to find out how useful web analytics is when used on various different types of websites. As found in this study, it is useful to use the data from web analytics to follow user pathways and highlight problematic areas for users – represented in error messages. However, when viewing

information based websites this type of data is likely have limited value in representing user experience. For information based website, comprehension of the messages on the site is critical for site success. This implies that the suggestions made of how web analytics may be used to heighten awareness of user experience on a website may be limited to transactional or gaming websites where user pathways are telling of their experience, however further research would clarify this proposition.

9 Recommendations

The results from this study have been very insightful in determining the strengths of both web analytics and think aloud when applied throughout the UCD and Evaluation phases of website projects. Having explored in detail where each method may be best used, the following list provides specific recommendations of their application, based on the findings of this study.

- During User Requirements analysis of a website redesign project, explore existing web analytics data to uncover usability issues which affect the current user population. If website owners are aware of problems on specific web pages through user feedback, or for example through poor web sales, explore the data which represents these pages in more detail. Additionally, web analytics data should be explored to determine where users are "dropping off" and whether there are usability issues which result in users leaving the site.

- Use web analytics data to base scenarios for think aloud evaluation on an existing website. The results of think aloud studies should be used to provide further understanding of issues highlighted by the web analytics data. Furthermore, take note of general comments made by users which may uncover further usability problems which were not highlighted in the analytics data, yet may be taken into account in new designs.

- Use think aloud evaluation to validate early design ideas, ranging from paper prototypes to fully functional wireframes.

- Implement all designs which take into account user requirements and user feedback gathered at earlier stages of the project.

- During Usability Evaluation run think aloud studies to gather further user feedback on final designs. If revising an existing live site, changes may be implemented

incrementally. Web analytics data which covers the time at which the new designs are in place may then be monitored to determine success of new designs.

- Launch the website. Following launch, use analytics data to determine whether user experience is improved quantitatively by exploring whether the issues found in the initial phases of redesign have been overcome.

- Use web analytics during Maintenance of the website to provide constant awareness of user pathways and potential usability issues. If issues are uncovered through web analytics, explore these further in the live data and use think aloud studies to uncover the reasoning from both sessions.

- Continue to iteratively test the site using analytics and where errors and drop off rates seem too high organise further think aloud sessions to explore the reason for such errors.

These clear recommendations are based on the findings of this study which were strengthened through the thorough and professional methodology used. Despite these strengths there were some limitations which may have limited the findings.

10 Reflection on methodology

The methodology used in this study was created to ensure the areas of interest including think aloud evaluation and web analytics could be explored in great detail. The following section firstly details validation of the results found in this study. This is followed by discussion of the limitations on the study, some of which were imposed because of the academic nature of the project with limited resources and some of which were simply uncontrollable. This section also highlights the strengths of the methodology, which ensured that valuable results and recommendations were made.

10.1 Validation

The results of all studies were briefly validated on completion by John Dumas, Director of Operations at optimum.web. This ensured that the findings from the tests were in line with a fellow usability practitioner. Ideally further validation of results would have been done by numerous usability professionals as well as web analytics experts to ensure that the findings were not affected by contamination because only one moderator completed all three phases of this study. In addition to contamination between studies, if further analysts were involved in

validating the results, potential problems caused by the evaluator effect (Hertzum and Jacobsen, 2001) would have been limited and potentially explored further.

10.2 Limitations

Despite gathering a vast range of insightful findings throughout this study, the following section details limitations of the methodology which were difficult to control.

10.2.1 Participants

The participants of both the controlled web analytics phase and the think aloud phase of this study were not truly representative of William Hill's user population. Although a number of the participants had some experience in either online or offline gambling, this experience was limited and unlikely to be typical of regular William Hill users. This limitation is likely to have generated results which may relate to participants' difficulty with basic concepts of online gambling, and therefore are unlikely to affect the true user population. Conversely, if William Hill hope to increase their user population to include less traditional gamblers by offering more mainstream events to bet upon such as casino games and television programmes, the terminology and methods used which were found to be problematic by the inexperienced participants of this study should be redesigns to ensure simplicity for all user types.

10.2.2 Knowledge of methods

Although I met with Andrew Stockwell for an induction to the RedEye analytics interface, this did not provide me with as much experience in analysing web analytics data as professional web log analysts. This inexperience may have reduced the amount of information I could infer from the web analytics data. Additionally, experienced web analytics analysts may be aware of certain methods which could be used to interpret the data more quickly, or infer more about user behaviour which I was not aware of.

Furthermore, I had little experience of the William Hill website. Although this was a positive factor, in that my previous experiences with the site did not affect any findings from all parts of this study – when exploring web analytics, which is based on codes which represent various pages of the website, thorough knowledge of the site may have been advantageous.

10.2.3 Contamination of findings

The order in which all three studies were completed was decided upon to reduce potential contamination of results. However, because all web analytics data had been analysed by myself prior to the think aloud studies taking place – also being led by myself, it was difficult to

completely ignore the findings already uncovered. This limitation may have resulted in my leading the think aloud studies to explore issues highlighted in the web analytics part of the study in more detail. However, without this prior knowledge of such issues they may have been seen as less serious and had less time spent on them during the think aloud sessions.

10.2.4 Controlled analytics study

The controlled web analytics part of this study did not help to influence the overall findings of the paper. The benefit of web analytics is the vast quantity of data gathered. However, the controlled exercise failed to gather a worthy quantity of data and therefore had limited value. Knowledge of users objectives was also fairly useless because for the types of tasks users were asked to complete, the pathways they took were obvious and could have been inferred without knowing their objectives e.g. Registration.

10.3 Strengths

Although the outlined limitations may have affected the value and purity of results, this study also had a number of strengths which generated a wealth of useful findings.

10.3.1 Support from commercial business

The support by optimum.web and Red Eye International were crucial to the success of this project. Particularly important was access to the live web analytics data provided by RedEye, which allowed me to experience analysis of data which represented the real user population. Without viewing this data it would have been impossible to gain such an insight into how usability issues may be highlighted by web analytics. Additionally, the process of analysing such a large amount of live data ensured that my experience was realistic to professional web analytics analysts, therefore highlighting the problems and limitations of analysing such data.

Further to access to real web analytics, support from optimum.web and RedEye provided me with information of realistic client demands and requirements from usability evaluation projects. Information from William Hill, detailing their specific interest in the registration process was important to remember during this project because it also informed me of the importance of such client objectives when determining how successful usability evaluation methods are.

10.3.2 Previous experience in think aloud evaluation

My two years of previous experience in running and analysing commercial think aloud evaluation studies have provided me with expertise in drawing conclusions from observation and analysis of user actions and comments. This experience helped to ensure that the think

aloud sessions derived as much information as possible from participants, which may not have been possible without previous knowledge of the process and analysis involved. Despite this strength, due to limited resources, the think aloud sessions were run and analysed by myself. Although this reduced the potential problems discussed by Norgaard and Hornbaek (2006) it also meant that the Evaluator effect discussed by Hertzum & Jacobsen (2001) was not explored.

Following discussion of the results found in this study, critical exploration of the methodology was useful to determine where the project excelled and also where improvements could have been made to further the study. Following this, a brief conclusion sums up the findings gained in this paper.

11 Conclusion

Despite the recommendation to combine web analytics and think aloud methods to gain the most realistic representation of user experience, every web usability project is different. This means that the methodology used must be adapted to ensure client objectives are matched, which will be specific to the type of website being evaluated as well as meeting time and budget limitations.

Web analytics is valuable at representing the true user population using a website in their natural environment. However, its known limitation of highlighting issues without insight into the reasoning for such issues devalues this method when used independently. This drawback of web analytics may be overcome through running think aloud sessions to gain first hand user feedback on all usability issues faced when using a website. Furthermore, results of think aloud studies may be validated by referring to web analytics data to ensure findings match the experience of the real user population overcoming issues of limited and unrealistic participants (Hong & Landay, 2001). Additionally, combining the results of think aloud with web analytics data reduces potential inconsistencies caused by the evaluator effect (Hertzum & Jacobsen, 2001).

This study explored the value of web analytics and think aloud methods in usability evaluation, to better understand user experience. In conclusion, this paper has confirmed a number of the benefits and drawbacks of both methods already set out in published literature. However in addition to this, recommendations have been made detailing how to combine both methods to overcome the known limitations and to therefore gain more accurate and detailed findings during usability evaluation. Furthermore, application of the methods throughout the User

Centred Design process is suggested to ensure that user experience is not only better understood but usability practitioners also take insights about user experience into account at all stages, truly keeping the end user at the centre of the design project.

12 References

12.1 Literature

Abras, C., Maloney-Krichmar, D., Preece, J. (2004) *User-Centered Design.* In Bainbridge, W. Encyclopedia of Human-Computer Interaction. Thousand Oaks: Sage Publications. Available from: http://www.ifsm.umbc.edu/~preece/Papers/User-centered_design_encyclopedia_chapter.pdf.

Addwise. (2001). *How to Gain Continuous Insight Into Your User's Problems: A New Era in Automatic Usability Testing.* White paper prepared by Addwise Infomange Ltd.

Arroyo, E., Selker, T., and Wei., W. (2006). *Usability Tool for Analysis of Web Designs Using Mouse Tracks.* CHI 2006, April 22-27, 2006. AC

Atterer, R., Wnuk, M., Schmidt, A. (2006). *Knowing the User's Every Move – User Activity Tracking for Website Usability Evaluation and Implicit Interaction.* International World Wide Web Conference Committee. WWW, 2006, May 23-26, 2006. ACM.

Chi, E. (2002). *Improving web usability through visualisation.* Internet Computing, IEEE.

Dix, A., Finlay, J., Abowd, G., Beale, R. (2004). *Human-Computer Interaction. Third Edition.* Prentice Education Limited 2004.

Guzdial, M., Santos, P.J., Badre, A., Hudson, S.E., Gray, M.H. (1994). *Analyzing and Visualizing Log Files: A Computational Science of Usability - all 9 versions.* GVU Centre TR GIT-GVU-94-8, Georgia Institute of Technology, 1994

Hackos, J. & Redish, J. (1998). *User and Task Analysis for interface design.* Wiley, USA.

Hertzum & Jacobsen (2001). *The Evaluator Effect: a chilling fact about usability evaluation methods.* International Journal of HCI, Vol 13, no 4, pp 41-443.

Hong, J., & Landay, J. (2001). *WebQuilt: A Framework for Capturing and Visualizing the Web Experience.* WWW10, May 1-5. Hong Kong, ACM.

Johansen, S and Hansen, J. (2006). *Do We Need Eye Trackers to Tell Where People Look?.* CHI 2006, April 22-27. ACM.

Mc Mullen, S. (2001). *Usability testing in a library web site redesign project.* Reference Services Review, 2001 - digitalcommons.rwu.edu

Nielsen, J. (1993). *Usability Engineering.* Academic Press, Boston, MA.

Nielsen, J., Clemmensen, T., Yssing, C. (2002a). *Getting access to what goes on in people's heads? – Reflections on the think-aloud technique.* NordiCHI. October, 19-23. ACM.

Norgard, M and Hornbaek, K (2006). What do usability evaluators do in practice? An explorative study of think-aloud testing. *DIS 2006.* Pennsylvania, USA. ACM Press.

Norguet, J.P., Zimanyi, E. & Steinberger, R. (2006). *Improving web sites with web usage mining, web content mining, and semantic analysis.* In Proc. of the 32nd Int. Conf. on Current Trends in Theory and Practice of Computer Science, SOFSEM. Springer-Verlag, 2006.

Peterson, E. (2007). *Global Use and Attitudes - Web Analytics Survey, March 2007.* Web Analytics Demystified Research Covering Global Use and Attitudes about Web Analytics Technology and Processes. Available from:
http://www.webanalyticsdemystified.com/link_list.asp?I=White%20Paper

Preece, J. (1994). *Human-computer Interaction.* Addison-Wesley, England.

Preece, J., Rogers, Y., Sharp, H. (2002) *Interaction Design*, *beyond human-computer interaction*. John Wiley and Sons, Inc. New York.

Ramey, J., Boren, T., Cuddihy, E., Dumas, J., Guan, Z., Van den Haak, M & De Jong, M. (2006). *Does Think Aloud Work? How Do We Know?* Conference on Human Factors in Computing Systems. ACM.

Rimmer., J, Wakeman,l., Sheeran,L & Sasse, M.A. (2000). *Messages from a Tangled Web.* Proceedings of OZCHI'2000, Sydney, Dec. 2000. pp. 4-11. Available from: www-mice.cs.ucl.ac.uk

Rosenfeld, L & Morville, P. (2002). *Information Architecture for the World Wide Web.* 2nd Edition. O'Reilly, CA, USA.

Rudd, J., Stern, K., Isensee, S. (1996). *Fidelity prototyping debate.* Interactions. January 1996, pp 77-85.

Scanlon, T. (1998). *Paper Prototypes: Still Our Favourite.* User Interface Engineering. Originally published: May 01, 1998. Available from: http//www.uie.com/articles/paper_prot otyping.

Shneiderman, B (1998). *Designing the user interface. Strategies for effective Human-Computer Interaction.* Third Edition. Addison Wesley Longman Inc.

Snyder, C. (2003). *Paper Prototyping – The fast and easy way to design and refine user interfaces.* Published by, Morgan Kaufmann.

Spiliopoulou, M. (2000). *Web Usage Mining for Web Site Evaluation – making a site better fit its users.* Communications of the ACM, vol 43, no. 8.

Spool, J. & Schroeder, W. (2001). *Testing Web Sites: Five Users is Nowhere Near Enough.* In Proceedings of ACM CHI Conference on Human Factors in Computing Systems. Seattle, WA.

Sterne, J. (2006). *Web Metrics Versus Web Analytics.* E-metrics article. Available from: http://www.emetrics.org/articles/

Usability and Accessibility Center (2007). *Approach – User Centered Design.* Image available from: http://usability.msu.edu/approach.asp

Weischedel, B. & Huizingh, K. (2006). *Website Optimization with Web Metrics: A Case Study.* ICEC, August 2006.

Wixon, D (2003). *Evaluating Usability Methods; why the current literature fails the practitioner.* Interactions, July – August 2003.

Wood, F., Siegel, E., LaCroix, E., Lyon, B., Benson, D., Cid, V., Fariss, S. (2003). *A Practical Approach to E-Government Web Evaluation.* IT Pro, May-June 2003. Published by the IEEE Computer Society.

12.2 Meetings

Dumas, J. (2007). Meeting with John Dumas, Director of Operations, optimum.web, 10 June 2007.

Stockwell, A. (2007). Meeting with Andrew Stockwell, Account Director, RedEye International. 6 June, 2007.

13 Appendices

Appendix 1 – Ethical consent form

UNIVERSITY COLLEGE LONDON

MSc & GRADUATE DIPLOMA Human-Computer Interaction with Ergonomics

CONSENT FORM

Project: *How web analytics may be used together with think-aloud studies in web evaluation, to further understand user experience.*

Name of the researcher: *Elizabeth Atkinson*

Please read and tick the following boxes if you agree:

1. I confirm that I have read and understood the information sheet for the above study and have had the opportunity to ask any questions and have received sufficient answers. ☐
2. I understand that my participation is voluntary and that I am free to withdraw at any time, without giving any reason, without my legal rights being affected. ☐
3. I agree for the researcher to video record my self during the experiment. ☐
4. I understand that any data collected during the session will be kept anonymous and confidential. ☐
5. I agree to take part in the above study. ☐

Use of data

6. I agree for the researcher to use pictures and video of me during the experiment to be included in the thesis. ☐
7. I agree for the researcher to use pictures and video of me during the experiment to be included in articles to be published in academic journals, conference proceeding and other equivalent articles. ☐
8. I agree for the researcher to use pictures and video of me during the experiment to be presented in conferences and other equivalent academic presentations ☐

_____   _____   _____
Participant's name   Signature   Date

ELIZABETH ATKINSON
_____   _____   _____
Researcher's name   Signature   Date

Appendix 2 - Web analytics instructions

Instructions

Imagine you are interested in betting on a couple of future sporting events. You usually place bets in betting shops, however you have heard that you can now bet online, so you decide to use the William Hill website to place your bets.

TASK 1) Go to www.williamhill.co.uk

TASK 2) Register with the William Hill website. Please do not proceed unless you see a confirmation of your Account number.

NB. During the registration process you will be asked to choose a user name.

Rebecca, your user name is: owltestingWA1

If you do not wish to use your real personal details for the remaining parts of registration, please use fake names/addresses – but please try to make them as realistic as possible or use the following:
E-mail address: asdf1@asdf.co.uk
Postcode: SW1W 0GT
Address: 26 Grosvenor Gardens
Telephone number: 0208 000 0000
Mother's Maiden name: Smith

TASK 3) Having registered, please place a bet on England to win the 2009 Ashes Cricket tournament, against Australia (do not enter a stake in GBP, please leave this box empty and simply add the bet to your bet slip).

TASK 4) Next, please place a bet on USA to win the 37th Golf Ryder Cup (do not enter a stake in GBP, please leave this box empty and simply add the bet to your bet slip).

TASK 5) Having added these bets to your bet slip, please now log out.

TASK 6) When you see a confirmation of logging-out, please now view your bet slip so that you can see both the Cricket and the Golf bets.

TASK 7) When you can see both bets, please clear your bet slip and log out of the website again.

Thanks, that's it!

Lastly though, please can you fill in the following details, and email it back to me.

Date of completion:…………………………………………………………………………………
Time of completion:…………………………………………………………………………………
Account number: …………………………………………………………………………………

Appendix 3 - Think aloud task script

optimum·web
usability & accessibility specialists

William HILL

Usability Test 19th July 2007

Pre-session – Moderator notes

➢ Points of investigation do not necessarily include all points of enquiry; they are for session guidance only, highlighting the key areas of investigation.

➢ Sessions to be conducted at 1024x768 screen resolution.

➢ Create/copy Favorites/shortcuts to the test PC desktop:
< www.williamhill.co.uk>.

Initial discussion with user: circa 5 minutes

Moderator notes – opening discussion with the user:

➢ Ensure users consents to testing by signing the ethical consent form.
➢ Explain to the user session format, scenarios, think-aloud, equipment, etc.
➢ Reassure the user that they are not being tested, they are helping us test new pieces of design for a web site.
➢ Briefly discuss with the user their situation, to include what they do and what their role is, and how they may use the Internet in their life – mainly non-work.
➢ Remind the user that they are able to withdraw from the test at any point.

Moderating the user session

Moderator style throughout the session:
➢ Throughout this session moderators should ask non-leading questions to draw out the user's reactions and thinking
             e.g. what are you looking at? what are you thinking?
➢ The moderator also has to draw out the user's expectations of what would happen, when the user clicks on non-working links and buttons
➢ The moderator should provide only the minimum input – taking care to avoid influencing the user's responses
➢ If users cannot achieve something, reassure them that it is not their fault, and the aim of the testing is to help improve the design
➢ Avoid giving help except where essential
➢ In debriefs ask open, non-leading questions...   ...and wait for (full) responses.

Objectives of the testing

Overall objective: To test the ease with which respondents move through the William Hill website, and are able to complete given tasks.

Scenario

Imagine you are interested in betting on a couple of future sporting events. You usually place bets in betting shops, however you have heard that you can now bet online, so you decide to use the William Hill website to place your bets.

*Firstly please click on this link **[www.williamhill.co.uk].***

Exploration of the William Hill home page

- Do users comment on their first impressions of the site?
- Is users' attention drawn to any particular area of the site?
- Do users notice the registration promotional panel?
- Do users independently explore any area of the site?

Now, please attempt to register with the William Hill site.
- Do users know how to register?
- Do they notice the create account link?
- Do users attempt to enter a username prior to registering for the site?
- Are users surprised by anything regarding the registration links?

Registration Process
- Are there any problems with the registration process?
- Do users notice the required fields?
- Do users notice all instructions around criteria for entered data?
- Do users have any specific problems when entering their personal details?
- Are all labels for fields clear, and do users know what they are expected to enter?
- Do users understand al terminology e.g. maximum bet?
- Do users know how to activate data entry fields e.g. drop down boxes?
- Do people have any problems with entering their security details?
- Do users realize they must accept terms and conditions?
- Do users understand the security question field?
- Are users familiar with the 'security question'?
- Do users find the 'security question' useful and does it make them feel that the account will be more secure?
- Do users comment on the amount of personal details to be added?
- Do users read the criteria of what the username/password field must contain?

Now you have registered, please place a multi-bet on England winning the 2009 Ashes Cricket tournament.
- Do users know how to add a bet to their bet slip?
- How do users go about doing this?
- Do users easily locate the cricket section?

Cricket Section
- Which link do they use to enter the Cricket section?
- In the Cricket section, do users find the information of the 2009 Ashes tournament?

- Do users understand how to place a bet on this tournament?
- Do users understand the 'Home Vs Away' terminology?
- Do users understand how to select the team they wish to bet on?

Now you have registered please place a multi-bet on England to win the 2010 Football world cup….

- Do users find the Football link?
- Which link is used to access the football section?

Football Section
- Is the 'add further selections' button used from the Betslip?
- In the Football section, do users find the information of the 2010 Football world cup?
- Do users comment on the length of the page within the Football section?
- Do users understand how to place a bet on this tournament?

Now please clear the bet slip and leave the website as you usually would.

Thank you.

General debrief

- **How was that for you?**

- **How easy was it to understand what you were required to do?**

- **Any specific areas for improvement?**

- **Anything you particularly liked about any designs in particular, or a combination of designs?**

- **Anything in particular stand out for you, bad or good?**

Thank you
Your participation today has been very helpful.