DOCTOR OF PHILOSOPHY

Consistency of Human Information Behaviour and its Impact on Personalised Contextualised Information Retrieval Systems

Grzywaczewski, Adam

Award date:
2013

Awarding institution:
Coventry University

Consistency of Human Information Behaviour and its Impact on Personalised Contextualised Information Retrieval Systems

Adam Grzywaczewski

Faculty of Engineering and Computing

Coventry University

A thesis submitted in partial fulfilment of the University's requirements for the Degree of

PhilosophiæDoctor (PhD)

January 2013

Thesis funded by EPSRC CASE Studentship in collaboration with Trinity Expert Systems Limited

1. Reviewer:

2. Reviewer:

Day of the defence:

Signature from head of PhD committee:

Abstract

The amount of digital information available in the Internet and various Intranets often causes information-overload, significantly increasing the amount of time and cognitive resources needed to acquire relevant and accurate information. When searching for information to address complex problems, users spend significant amount of time clicking through search-query results and reformulating the search if the results are not satisfactory. This is a tedious and challenging task and has a negative impact on the global economy.

Personalisation and contextualisation techniques intend to address the above mentioned problem. Such techniques help to derive additional information from the history of the past user interactions (user profile) or the current context of interaction. This information is used to refine the search results in order to narrow down the scope of search queries for better results.

One of the key requirements for the development of personalised or contextualised search utilities is consistency of human behaviour. It is only possible to predict user future preferences and actions if they are correlated with past behaviour. This fact is frequently ignored in the current IR literature even though empirical evidence clearly illustrates that humans are very inconsistent when interacting with information. This leads to very low predictive validity of existing contextualisation/personalisation IR implementations.

This thesis hypothesises that HIB could be consistent under certain contextual and task specific conditions. The thesis claims that a large proportion of our daily information activities are highly consistent and also meet the definition of habitual behaviours. In other words, even though empirical evidence clearly illustrates that on average human information behaviour is

very complex and dependent on thousands of factors, there will exist a small group of daily activities which are highly consistent and can be supported effectively by modern IR utilities. As a consequence the development of highly effective personalised or contextualised solutions are feasible.

In order to prove the above mentioned hypothesis User Study 1 (diary study) was carried out. User Study 1 revealed that a significant proportion of our daily information interaction is indeed consistent (49%) and a significant proportion of this can be classified as habitual (41.9%). User Study 1 also confirmed that the behaviour of participants is consistent only when the same tasks are carried out in the same context and under the same emotional user state. Finally User Study 1 confirmed that the behavioural consistencies are highly individual and affected by a number of external factors. However there exists no HIB model or research methodology which can help to identify the external factors systematically and take advantage of them during IR system design. Therefore this research proposes an "Integrated Framework for HIB in-situ" which intends to fill the above mentioned gap in knowledge.

The framework was designed to support the development of Information Retrieval utilities based on consistent behaviour of clearly defined user groups and their problems. In order to illustrate its applicability a single user group was selected. The proposed framework was successfully applied to the problem of work-related activities of software engineers. The framework allowed for identification of a more specific user group of software developers and narrowed down the investigated task to code development/debugging. In the next step the framework allowed for shortlisting a number of behaviours which had a significant potential for consistency. The consistency of the shortlisted behaviours and their correlation to relevance was verified through User Study 2 (questionnaire). The key shortlisted behaviours were further analysed through User Study 3 (fully automated, long lasting ethnographic study) which allowed for the identification of factors that have a key impact on behavioural variance. The analysis revealed a number of consistent behaviours (implicit-feedback indicators) that can be used for prediction of document relevance. Importantly User Studies 1 and 3 validated the research hypothesis on a specific case study of software engineers. However the proposed framework is based on very basic cognitive mechanisms responsible for the human decision-making process and as a consequence is highly generalizable to other user and problem groups.

To my parents, for preparing me for this challenge

And to my wife, for supporting me throughout

Acknowledgements

I would first like to acknowledge the companies and organisations that supported me during my PhD and especially Trinity Expert Systems Limited (TESL), the Engineering and Physical Science Research Council (EPSRC) and Coventry University and Coventry University Business Development Support Office (BDSO). This journey would not be possible without the contribution of countless people from those organisations and their financial and direct support.

My supervisor, Dr Rahat Iqbal and my supervisory team members, Dr John Halloran and Professor Anne James, have all had a profound impact not only on the direction and shape of this work but also on my overall graduate experience. Their support is not possible to quantify. Together they have protected me intellectually, administratively and financially.

Dr. Iqbal has been a mentor to me throughout my Ph.D. and regardless of my sometimes ridiculous working hours he was always able to find time for me and make me a priority. I would like to thank him for countless hours spent on revising my ideas and helping me to improve my very difficult (to say the least) writing style. Dr. Iqbal taught me how to step back and look at my work from the perspective of other people. Without the patience, empathy and perseverance that he demonstrated everyday, the work on this PhD would not have been possible.

Also, retrospectively, I would like to thank him for starting this journey by offering me a PhD position, believing in me more than I have believed in myself, and suggesting the application for the EPSRC CASE Studentship and guiding me and the collaborating company through the application and submission process.

I would also like to thank him for engaging me in a countless number of other funding and teaching projects. The opportunity to participate in such a wide range of academic activities opened my eyes on the world of possibilities and prepared me for future academic work.

I would like to thank James Butler, Head of Enterprise Technology of Trinity Expert Systems Limited, for believing in me and my ideas. Without him and his direct support this project would not be possible.

Similarly I would like to thank Chris Drinan, Coventry Development Team Manager of Trinity Expert Systems, for his flexibility in managing my time and open-mindedness in my project allocation.

I am aware that there are too many people to mention personally therefore I would like to thank everyone who throughout my PhD supported my work directly or through revisions, comments or criticism.

I would also like to thank my family and closest friends for tolerating my late nights and never-ending lack of availability, especially my wife, Izabela Grzywaczewska, for taking a significant proportion of my responsibilities and factoring my PhD into all of our plans which was not always easy.

Finally, I would like to thank my parents for supporting me emotionally throughout the process but also for shaping me the way I am now. Thank you for encouraging me even in the most difficult moments. The majority of my success can be attributed to you.

I would like also to thank all of the participants who took part in my studies and the people that supported me with their skills of organisation and administration.

Contents

List of Figures

List of Tables

Glossary

Attitude (Psychology) A psychological tendency that is expressed by evaluating a particular entity with some degree of favour or disfavour (Eagly and Chaiken, 1998).

Attitude Toward the Behavior Is a degree to which the behaviour is positively or negatively perceived.

Black Box Model An approach to the analysis of complex systems without the need to understand the construct of its internals.

Human Computer Information Retrieval (HCIR) A research area in library and information science that lies on the intersection of Human-computer interaction and Information Retrieval (IR).

Human Information Behaviour (HIB) A research area in library and information science that focuses on investigation of human interaction with information resources. The discipline lies on the intersection of library and information sciences, cognitive psychology and Human Computer Behaviour.

Inforamtion Relevance Is a subjective value denoting the extent to which an information meets the Information Need of a user.

Information Litearcy Ability to make effective use of information sources including libraries, catalogue systems, computer systems and other sources.

Information Need Is a desire, experienced by an individual or a group, to obtain an information.

Information Object An entity of information with which the user or computer system interacts.

Information Overload A state at which an individual is exposed of too much information.

Information Poverty A state at which an individual does not have the information that one requires nor the capability to find it.

Information Retrieval (IR) A research area in library and information science that focuses on investigation of the process of finding relevant information in large information systems and design of such systems.

Information Seeking and Retrieval (IS&R) A research area in library and information science that focuses on investigation of the process of finding relevant information regardless of the location and format of the retrieved information.

Modearting Factor A factor that affects the relation between two other variables.

Perceived Behavioral Control The amount the control the human believes he has over a behaviour.

Predictive Validity Is a measure of agreement between results obtained by an indirect instrument (e.g. implicit relevance-feedback evaluation) and results obtained through a more direct observation.

Subjective Norm Represents a perceived social pressure related to a behaviour.

Systems Oriented approach to IR An approach to Information Retrieval that focuses on the design of information retrieval systems marginalising the role of the user.

User Oriented approach to IR An approach to Information Retrieval that focuses on the user interaction with information retrieval systems.

1

Introduction

1.1 Introduction

The importance of Information Management and Information Retrieval systems is unquestionable in the modern society and both individuals as well as companies recognise the benefits of being able to find information effectively. In contrast to the basic Internet search dominated by Microsoft, Google, etc., the specialist or enterprise search solutions are still underdeveloped. 'Page Rank' and similar algorithms which are very effective when working with highly cross linked HTML do not work:

- outside of the Internet domain (e.g within Intranet networks);
- with more complex data types;
- HTML not meeting the search engine requirements (e.g. not cross linked or dominated by non HTML content);
- when the HTML has a dynamically changing nature (e.g. Internet shops, mobile application stores, source code repositories, image repositories, etc.).

Without the back-link information such systems do not have any basis for assessing the relevance of the information apart from the direct comparison of the query to the textual content of the document. Achieving high levels of relevance based exclusively on semantic analysis, because of a complex nature of human language and the subjective nature of the concept of relevance, is very difficult (if possible at all). To make things even more complex for the majority of users formulating an effective search query, especially from outside of their domain expertise is a cognitively challenging task. The unique qualities of individual expression, limit keyword-based search engines and Information Retrieval (IR) systems.

As a consequence for years now there exists a high demand for dedicated Information Seeking and Management utilities, that are designed in order to support needs of dedicated user groups, that have to work with non Internet (e.g. Intranet) or non mainstream information.

This thesis is a direct result of the author's observation of this problem during his work as an Information System developer and then architect in both Microsoft Poland and Trinity Expert Systems Limited (partner and sponsor of this research). When working with multiple customers it was normal process to ask the information workers about properties that constitute valid and valuable information and to try to understand how they interact with, search and re-use that information. Such analysis led usually to the development of complex information systems integrating their everyday work tools (e.g. Outlook, Word, Excel and SAP) with Information Seeking and Management utilities custom designed for particular group needs. After a while it became apparent that even though people are having issues describing the properties of relevant information they are very good at identifying it once they have found it. This leads to the idea illustrated in Figure 1.1 which is based on the assumption that the users will be capable of identifying relevant information and providing feedback to the computer system. This feedback can be provided either explicitly, through users evaluating document relevance or implicitly through the system observing the behaviours that can indicate relevance (e.g. printing the document).

Feedback based personalisation and contextualization described above, is very promising and for years now has attracted a lot of attention in the Information Seeking and Retrieval (IS&R) domain. At the point of initial investigation and largely until today it was widely believed that through gaining an understanding of user intentions and the environment in which the user interacts, IR systems can be improved (Teevan et al., 2010). It is believed that the knowledge of accurate search context or the user profile can allow IR systems to improve the search experience even when ambiguous search queries are used (e.g. distinguish whether the user searching for a jaguar is interested in an animal; car marque; Mac OS version; British heavy metal band or other) or when the relevance cannot be estimated using 'Page Rank' like algorithms.

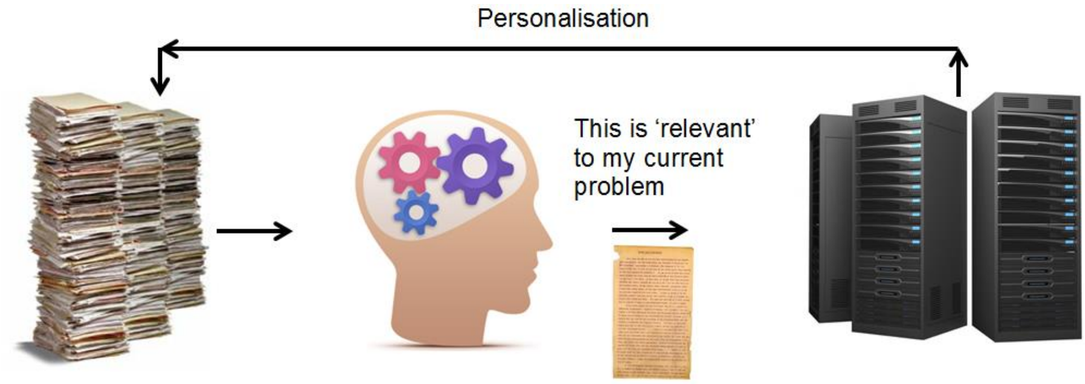


Figure 1.1: Relevance-Feedback IS&R System - High level diagram of the relevance-feedback IS&R System

Attractive as it may seem the design of universal relevance-feedback systems faces a number of challenges. Empirical results investigating the human perception of information relevance suggest that it is a very fragile concept, not only changing in time, but also highly affected by the widely misunderstood context of the interaction. What is more important, behaviours indicating relevance (e.g. implicit-feedback indicators such as printing or reading time) vary with a multitude of factors such as:

- Domain expertise (Kelly and Cool, 2002; Sihvonen and Vakkari, 2004; White et al., 2009; Wilson, 2000),
- Cognitive Ability (Tu et al., 2008; Xu and Wang, 2008),
- Task and Problem (Marchionini, 2006; Saracevic et al., 1997; Tu et al., 2008),
- Search expertise (Lazonder et al., 2000; Tu et al., 2008; White and Morris, 2007; Zhang et al., 2005),
- Presentation Order (Bar-Ilan et al., 2009; Keane et al., 2008; Purgailis Parker and Johnson, 1990; Xu and Wang, 2008),
- Personality (Halder et al., 2010; Heinstrm, 2003; Heinstrm, 2005),
- Cognitive Style (Ford et al., 2002; Graff et al., 2001; Kim, 2001*a*; Wood et al., 1996),

- Language and culture (Dong and Lee, 2008; Graff et al., 2001; Kralisch and Berendt, 2004; Li and Kirkup, 2007),
- Demographics (Almutairi, 2011; KiM, 2008*a*; Large et al., 2002; Roy and Chi, 2003),
- Information object (Buda and Zhang, 2000; Kelly and Teevan, 2003; Oard and Kim, 2001; Purgailis Parker and Johnson, 1990),
- Number of documents (Bar-Ilan et al., 2009; Huang and Wang, 2004; Murphy et al., 2006; Oulasvirta et al., 2009; Piorkowski et al., 2012*a*),
- Presentation format (Joachims et al., 2005; Keane et al., 2008),
- Presentation order (Elsenberg and Barr, 1988; Huang and Wang, 2004; Mantonakis et al., 2009; Purgailis Parker and Johnson, 1990; Xu and Wang, 2008),
- Epistemological Believes (Marchionini, 2006; Tu et al., 2008; White and Roth, 2009),
- Metacognitive Ability (Ford, 2004; Ford et al., 2005; Gross, 2005),
- Physical Environment (Case, 2007; Cool, 2001; Dervin, 1997; Johnson, 2000),
- and many more (for a comprehensive literature review of factors affecting human perception of relevance refer to Saracevic (2007)).

The above provides a very limited list of empirical studies investigating the impact of various environmental, situational or personal factors on Human Information Behaviour (HIB). The review of theoretical work in HIB and especially the analysis of the model of HIB reveals an even higher level of human behavioural complexity. Models or methodologies such as Dervin's Sense Making (Dervin, 1983); Wilson's Model (Wilson, 2000); Belkin's Anomalous State of Knowledge Belkin (1980) or Ellis's Model of Information Seeking Behaviour Ellis and Haugan (1997), add tens of factors affecting the decision-making process. Further literature reveals thousands of additional factors spread across hundreds of models making the complexity of the problem unmanageable (i.e. the literature review carried out by Fisher and Julien (2005) reveals 72 distinct models with thousands of factors).

This does not prove that inferring understanding about information relevance from human behaviour is not possible. On the contrary all of the above research is founded on the assumption of human mechanistic consistency and machine like repeatability. The above research simply identifies an existence of a variation in human behaviour which is caused by the differences in the environment, problem or user over time. The fact that the models assume the machine like nature of a human being does not mean that they make the design process of IR computer systems easier or provide utilities for managing the claimed behavioural complexity. On the contrary, the number of identified factors, complexity and the level of interaction between them makes the problem almost unmanageable.

The above had a significant impact on this research because it was not possible to systematically capture user domain expertise, cognitive ability and most of the factors listed above. Therefore it became necessary to seek other approaches and methodologies to support the system design and validation process. Even though the review of psychological literature which followed identified a number of potential approaches that could be used, it was not easily applicable within the IS&R domain. The literature review identified few cross references between IR and Library Science literature but more importantly even fewer references between widely understood IS&R research and psychological research. The above is widely recognised in the IS&R community and referred to in the literature as a 'Gap in Information Retrieval' (Belkin, 2008; Ingwersen, 1996; Saracevic, 1997*b*; Wilson, 2008)). White and McCain (1998) empirically illustrated the existence of the 'Gap' through numerous co-citation analysis studies (see the results of this research illustrated in figure 1.2). The existence of the 'Gap' leads to the situation where the multitude of the proposed models/meta theories is rarely supported with any empirical evidence (e.g. analysis of behaviour of real users, or design and validation of real IR systems) or the evidence is sparse and very generic.

Similarly, even though the cognitive mechanisms responsible for Information Seeking Behaviour are the same during problem-solving and "Information Foraging", there exists a very limited amount of literature referencing the findings of modern psychological or even Human Computer Interaction research (see figure 1.3).

The above is very unfortunate as the psychological and social research, throughout decades developed a number of methodologies and behavioural theories that can and indeed are successfully applied to the problem of prediction of human behaviour (i.e.

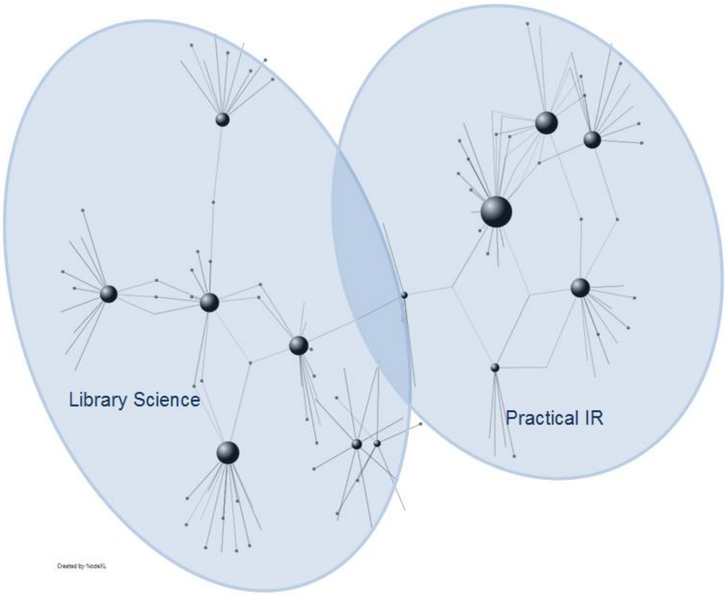


Figure 1.2: Co-citation IR vs Library Science - The gap in Information Retrieval illustrated through co-citation analysis

of health related behaviours, consumer behaviours, racist behaviours, voter behaviours and many more). Jackson (2005) reviews a wealth of theories (approximately 40) which are not only well founded in theoretical psychology research but also supported by experimental as well as real life/commercial data. Rational Choice Theory (Scott, 2000) and Adjusted Expectancy-Value Theories (Reynolds and Gutman, 1988), including the Theory of Planned Behaviour to which this thesis refers extensively (Ajzen, 2005) are two examples. Another example are approaches/methodologies referred to as Integrative Theories of Consumer Behaviour (Giddens, 1984). Even though frequently criticised for numerous limitations, has decades of track record and thousands of successful commercial and real life applications. Just the Theory of Planned Behaviour has been successfully applied to thousands of empirical studies effectively predicting and controlling human behaviour of various nature (Armitage and Conner, 2001).

It is also important to note that there is no reason why one should assume that HIB is in any way different from our everyday behaviour and interaction with non computer/library related information. It would be a highly risky assumption to claim that the cognitive processes involved in searching for information are different from those employed when solving other everyday problems, and indeed there exists no

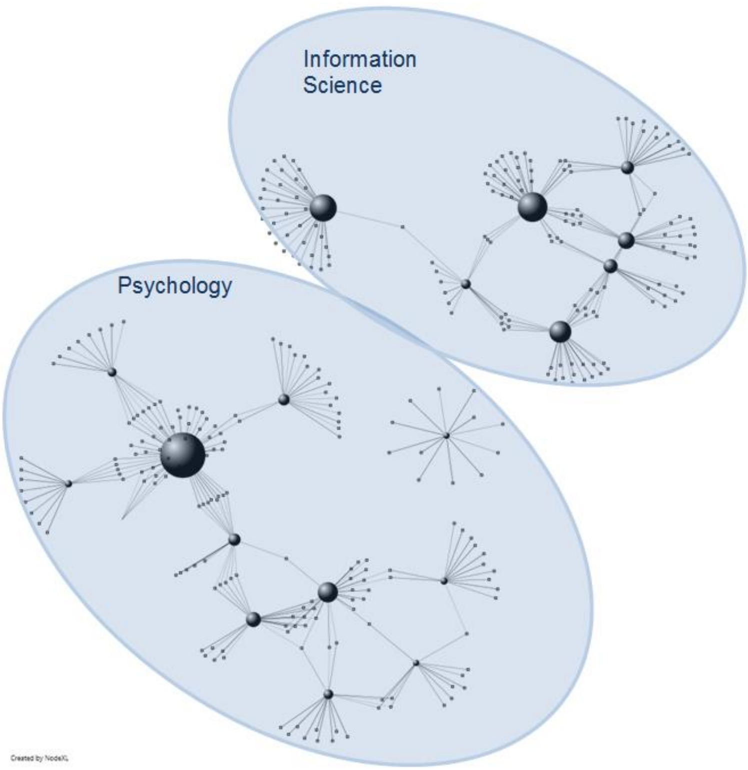


Figure 1.3: Co-citation Psychology vs Library Science - The gap in HIB illustrated through co-citation analysis

research trying to successfully claim that. On the contrary, there exists a branch of research investigating human information foraging Pirolli and Card (1999) which investigates the extent to which human beings apply their basic food gathering instincts when searching for information.

The review of the above mentioned literature and the evident gap in information sciences shaped the scope of this research. This study, which was born as an attempt to develop a highly personalised implicit-feedback search utility for software developers was transformed into a wider study of consistency/predictability of human information interaction. The lack of HIB specific literature investigating the fundamental concepts of human consistency led to the need to reinvestigate the concept of consistency and revisit the way it is being analysed in IS&R research. The above investigation led to the development of the proposed research framework ('Integrated Framework for HIB in-situ' discussed in chapter 4) and its application to the group of software developers (chapter 5). This also allowed the delivery of the primary goal of the EPSRC funding which was the design of the implicit-feedback system for software developer, work-related IR (chapter 6).

Currently, the results of investigation of HIB consistency and especially the philosophy of the proposed framework are being applied to a number of commercial research projects at Jaguar Land Rover Research as well as forming the foundation of further PhD studies. It is the author's ambition that with time and with further work invested in documentation it will be possible to transform what is now a research framework into a system design methodology/process. This will require the further documentation of research examples illustrating the process and further development of experiment tools, together with further dissemination not only in the research community but more importantly in the software development environment.

1.2 The problem of HIB consistency and the research hypothesis

An important aspect of the IR research is the assumption of human behavioural consistency. The problem of human behavioural consistency is not only central to this research but also a fundamental and key philosophical assumption of any research discipline. Personalised IR is only possible if there is a correlation between user past and future behaviour. It is not possible to consider any form of contextualisation or personalisation if Human Information Behaviour (HIB) is not at least partially consistent. Even if human behaviour cannot be perceived as mechanistic/functional a probabilistic model of human behaviour has to exist for the prediction of user information interests to be possible.

Despite the importance of the problem current research takes user stability and rationality for granted and rarely investigates the issues of Human Behavioural Consistency. The existence of a causal relationship between user past behaviour and future choices and needs became an axiom of many IS&R (Information Seeking and Retrieval) disciplines. This might be justified by the fact that people like to believe that their own thoughts, emotions and actions are rational and consistent within a given context. Unfortunately the evidence does not support this hypothesis. The empirical research (illustrated in Chapter 2) clearly shows that the human consistency is "more apparent than real" (Ajzen, 2005). The empirical research (Ellis, 1996; Kraus, 1995; Paunonen, 2003; Saracevic, 2007; Wicker, 1969) illustrates that the correlation between individual instances of human behaviour is very low and in most of the studies it rarely reaches p=.3 applying Pearson's correlation coefficient. Importantly, if humans are irrational and their individual behaviours are not correlated to each other it makes little sense to use past history of user behaviour to predict future user choices.

This thesis investigates the notion of human behavioural consistency in the context of IS&R. The goal of this thesis is to investigate the conditions in which Human Information Behaviour is consistent and personalisation/contextualisation of IR systems is achievable. The above goal can be summarised through the following research hypothesis:

> "Human Information Behaviour (under certain conditions identifiable by the proposed methodology) is consistent and as a consequence can be used for development of IS&R utilities"

In order to achieve the above the thesis investigates the psychological mechanisms (e.g. decision-making and learning processes) of human consistency and discusses how they impact the IS&R process. The thesis provides a detailed study of human general (psychological literature review) and information-related (HIB literature review) consistency and discusses the factors that affect the human variability. Furthermore it discusses how the variability of human behaviour affects the efficiency of the existing IR solutions and how the knowledge of factors affecting this variability can be used to improve the user search experience. The research proposes a new integrated framework that can be used in IS&R research and IR system implementation. The proposed framework can be systematically used to overcome the challenges introduced by user behavioural variability for effective IR design. Finally the proposed framework is applied to a case study of Software Developers and their work-related information problems in order to illustrate its validity and simplicity of use.

The rest of this chapter is organised as follows. Section 1.3 provides a background for the research. It discusses the overall problem of behavioural consistency and its importance for the development of personalised and context sensitive IR solutions. Section 1.4 discusses the motivation for this research. Section 1.5 illustrates the goal of the research and lists the key aims and objectives of the thesis. Section 1.6 introduces the research methodology listing the key steps required to achieve the highlighted goals. Section 1.7 briefly discusses the contribution of the research. Section 1.8 provides an overview of the rest of the thesis.

1.3 Background

Personalisation and contextualisation of the IR utilities is one of the key trends in IS&R research. There exists a significant body of research, both purely academic as well commercially driven (by Microsoft, Yahoo, Google and many other intranet search providers) focusing on the development of various personalised and contextualised IR solutions.

One of the most interesting branches of the new IR research focuses on interpretation of human interaction with computer systems to learn about user information interests and to tailor future search based on that knowledge. Such approaches interpret user micro-behaviour, such as reading time or the click pattern of the user, to estimate the relevance of the currently viewed documents (Kelly and Teevan, 2003). Through their behaviour users provide the IR system with implicit relevance-feedback which in contrast to the explicit feedback does not require the user to take any additional actions such as rating the content (Back and Oppenheim, 2001).

Many IR solutions use a rich combination of observable factors of human behaviour to decide about user perception of relevance, create a user profile or estimate the extent of user expertise (Hembrooke et al., 2005; Joachims et al., 2005, 2007; Kelly and Belkin, 2004; Kelly and Teevan, 2003; Lee et al., 2008; Poblete and Baeza-Yates, 2008; Spink and Losee, 1996; White and Kelly, 2006). The systems use the relevance-feedback to adjust the search results or the details of UI (displaying various forms of information depending on the IR context, e.g. map view, images, Wikipedia view) enhancing user experience or maximising advertising effectiveness.

The main criticism of the contextualisation and personalisation solutions (as discussed above) lies in the following assumptions to make the approach feasible:

- Human beings are consistent and rational.
- This rationality leads to the existence of correlations between elements of context and user behaviour.
- The correlation between the observable elements of context explains a significant proportion of behavioural variance.
- The correlation changes predictably in time.
- The correlation is not unique for individuals but can be generalised to wider user populations.

As one can see the notion of human rationality and consistency, its existence and extent, is a key for successful contextualised IR design. Unfortunately there exists still a very limited amount of research focusing on the problem of human information consistency during IR. As a consequence the extent of human information-related consistency and the way it is affected by moderating-factors is still not fully understood.

Even though the recent decade brought significant progress in our understanding of the mechanisms of user information-related interaction it also revealed a number of problems related to the behaviour based approaches to IR. The key problem faced by all of the personalised and contextualised IR approaches is a very high level of observed behavioural inconsistency that affects the performance. The observed HIB inconsistency results in low predictive validity of the solutions based on user behaviour prediction and as a consequence prevents the widespread commercialization of such utilities.

The empirical research on HIB has identified an inconsistency of user behaviour across individual instances of information interaction. The experiments on the use of click-through information, the utility of reading time as a source of relevance information or on the use of session level variables of interaction, confirmed the findings of theoretical models of HIB identifying a very high variability of human responses. The experimental results confirmed that humans do not make the relevance decision exclusively based on the search problem and the semantic similarity of retrieved documents. Instead the decision is highly dependent on an overwhelming number of factors related to the interaction context, task attributes and the character, experience and disposition of the user. For example:

- User reading time varies between the interactions depending on the extent of user domain knowledge, search engine expertise, task properties and cognitive ability. This makes the use of reading time as an implicit-feedback indicator very difficult, as the correlation of this variable with relevance, changes across different contextual settings, tasks and users (Case, 2007; Dervin and Clark, 1999; Dervin et al., 2003; Fidel et al., 2004; Fisher and Julien, 2005; Ingwersen and Järvelin, 2007; Saracevic, 1996, 1997*a*; Savolainen, 1995; Spink and Cole, 2005; Spink and Heinström, 2011; Wilson, 1999).

- User click through pattern is highly affected by the information presentation format, number of presented documents, user search experience, type and trust of the search tool and importantly the presentation order itself (Bar-Ilan et al., 2009; Elsenberg and Barr, 1988; Huang and Wang, 2004; Joachims et al., 2005, 2007; Keane et al., 2008; Murphy et al., 2006; Purgailis Parker and Johnson, 1990; Xu and Wang, 2008).

- The use of retention actions (e.g. copy and pasting, saving, bookmarking, printing, etc.) is affected by the nature of the task carried out and can change from zero to tens or hundreds daily (Cooper and Chen, 2001; Kim et al., 2004; Oard and Kim, 2001; Zhang and Seo, 2001).

The empirical research discussed above and extended in the literature review (chapter 2), clearly illustrates that HIB changes very rapidly along with the multiple contextual, user and task related factors, making the personalised/contextualised IR development infeasible until the nature of those variations is further understood.

This observation has theoretical support as well. The analysis of just two HIB models, presented in Figure 1.4, makes it clear that the human information response is dependant not only on the semantic content of the document, but also on a big amount of user, context and task related factors. Even though there are substantial differences between the Wilson's and Dervin's model they both clearly show that:

- The decision about the information relevance is based on a variety of factors.
- The way this decision is made and the impact of individual factors on this decision is a continuously changing process.

In the case of Wilson's model (Figure 1.4 left) the IR process is enclosed in a positive feedback loop. In this case the results of IR feed back to the person's perception of the problem and the context of retrieval. In the case of Dervin's model (Figure 1.4 right) the user is in a never ending IR journey travelling from one situation to another and never carrying out the IR process from the same perspective.

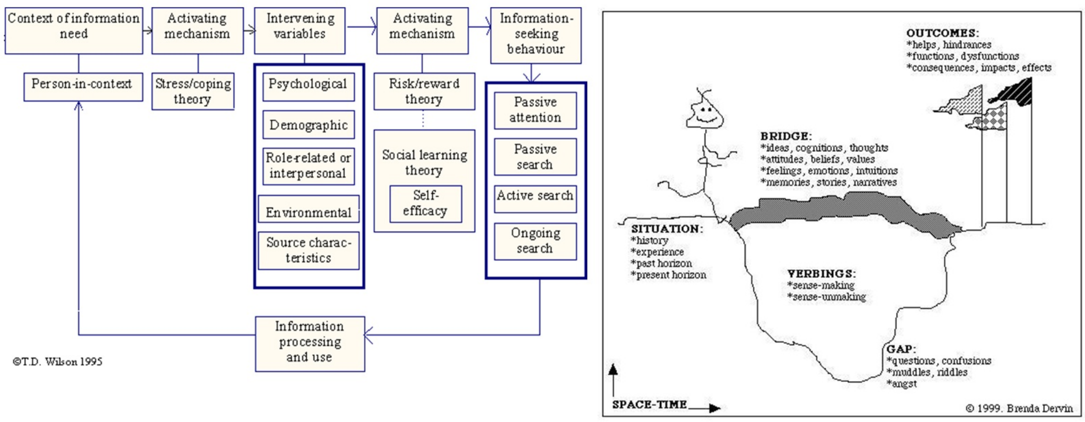


Figure 1.4: Two exemplary HIB models - Wilson's information behaviour model on the left and Dervin's sense making model on the right

The complexity of the information-related decision-making process, moderating-factors, their interdependence and frequency of behavioural change, makes the interpretation of user responses a very challenging task. As a consequence the development of personalised IR solutions based on behavioural analysis is complex and there is still a potential for further research in this domain. Such complexity, which manifests itself as an in-consistency of HIB across individual information interactions, raises questions regarding the extent of human rationality, consistency and as a consequence, predictability.

Even though a number of research methodologies exist intending to address the problem of HIB observed inconsistency (Case, 2007; Dervin et al., 2003; Fidel et al., 2004; Fisher and Julien, 2005; Ingwersen and Järvelin, 2007; Saracevic, 1996, 1997*a*; Savolainen, 1995; Spink and Cole, 2005; Spink and Heinström, 2011; Wilson, 1999), so far there does not exist a methodology that would overcome the challenges of HIB variability during the personalisation/contextualisation of IR.

The observed human inconsistency is not a unique challenge of IS&R research and various attempts have been carried out previously in the psychological domain in order to understand the human behavioural variability. The literature review in this domain is much more mature (i.e. more than 100 years of research) and as a consequence reveals even more the complex nature of the problem, but also offers a variety of ways to address it that can be reused in the IS&R domain. The research in Cognitive Psychology (Ajzen, 2005) confirms that the outcome of the problem-solving and decision-making process cannot be estimated in isolation nor through analysis of the limited number of moderating-factors (e.g. domain or web expertise). Moreover, the empirical research illustrates, that the predictive validity of approaches based on moderating-factors is very low and their correlation with the results of user action is rarely higher than p=0.3. This does not mean that the HIB is not rational or consistent but explains the low efficiency of the current IR utilities.

There exists a number of behavioural change theories explaining human behaviour with much higher accuracy. For example the "Theory of Planned Behaviour" explains more than 30% of behavioural variance. The theory is widely used for prediction of consumer behaviour, health or environment related behaviours (Ajzen, 2011; Armitage and Conner, 2001; Fishbein and Ajzen, 2009; Schulze and Wittmann, 2003). Similarly, theoretical and empirical research in habit development and behavioural automaticity suggests that a big proportion of our daily activities is highly repeatable (Ouellette and Wood, 1998; Wood et al., 2002). What is more important, the research results suggest that frequent and consistent past behaviour is a good predictor of future behaviour (Ajzen, 2002; Bargh, 2007) which supports the hypothesis of this research.

The results of the last 60 years of cognitive psychology research suggest not only a high feasibility of personalised and task sensitive IR development but also point out the difficulties of the existing approaches, especially approaches based on moderating-variables.

1.4 Motivation

The work presented in this thesis is motivated by the two key shortcomings of the existing knowledge:

- Lack of models/frameworks allowing for better understanding of the HIB consistency: Current HIB models provide us with an in-depth understanding of what environmental and user variables are impacting the HIB. They provide an extensive list of groups of factors that should be considered as part of IR context. The HIB research explains extensively the cognitive mechanisms responsible for IR related problem-solving and decision-making. Models/research methodologies such as the Sense Making Approach (Dervin, 1983; Dervin et al., 2003), the Holistic Cognitive Theory of IR (Ingwersen and Järvelin, 2007) or Wilson (2000) model of HIB provide the research community with a foundation for the successful design of IR and HIB experiments. This unfortunately does not have a direct impact on our understanding of human consistency/rationality which is a necessary condition for personalised/contextualised IR development. The majority of the existing models merely list the types of factors that affect HIB. They neither investigate the way the factors impact the behaviour nor discuss the way this impact can be captured, predicted or modelled. The work presented in this thesis intends to address the above by investigating the nature of the above mentioned effects and their repeatability/consistency.

- Lack of processes/methodologies that would allow for taking advantage of HIB consistency for the personalised/contextualised IR development: Current HIB as well as IS&R research also lacks frameworks/methodologies that can be used in the systematic HIB research/IR design process. As discussed in the point above, there exists a gap between understanding the factors affecting the HIB and the application of this knowledge for IR system design. This work intends to build on the knowledge about HIB consistency and propose a research framework that will help to fill this gap. The proposed "Integrated Framework for HIB in-situ" does not offer a universal understanding of HIB and all of the factors affecting it but it does allow for systematic research leading to commercially viable IR implementations.

1.5 Aims and objectives

The main aim of this research is to investigate when and to what extent, HIB is consistent and how the knowledge about human consistency can be employed to the design of effective IR solutions. To achieve the aim the following objectives have been set:

1. To investigate the existing personalised/contextualised Information Retrieval research in order to understand the HIB Consistency problem.

2. To investigate the factors impacting the Human Behavioural Consistency and as a consequence the efficiency of personalised/contextualised IR.

3. To investigate the feasibility of IR personalisation/contextualisation in conditions of low HIB consistency.

4. To validate the hypothesis of user consistency through user studies.

5. To propose a framework for contextualised/personalised IR system development overcoming the challenges of HIB variability.

6. To evaluate the applicability of the proposed framework through:

    - Applying the proposed framework to the case study in order to understand the nature of environmental/contextual consistency of the selected population.

    - Applying the proposed framework to the case study in order to understand the nature of behavioural consistency of the selected population for the identified consistent environments/contexts.

- Applying the proposed framework to the case study in order to identify the consistent behaviours that can be used for the development of personalised IR systems.

1.6 Research methodology

Figure 1.5 shows an adopted methodology consisting of 6 stages as described below:

- Stage 1: The research opens with a general investigation of the existing approaches to personalised and task sensitive IR. The review focuses both on the theoretical HIB research as well as the empirical IS&R literature.

- Stage 2: The second step relates to a much more focused literature review concentrating primarily on the problem of HIB and its consistency. The branches presented on the left of Figure 6 relate to the IR and Library Science research whereas the branches on the right focus exclusively on the Cognitive research. The goal of the review is to lay the foundation for the proposed generalised research methodology referred in the rest of the thesis as the "Integrated Framework for HIB in-situ".

- Stage 3: Based on the literature review a new methodology is proposed at this stage which addresses the limitations of the existing approaches. The methodology is applied at this stage in order to design the further user studies supporting the aims and goals of the research with qualitative evidence.

- Stage 4: Focuses on the investigation of the environmental/contextual consistency. The goal of this step is to investigate a wide range of user information activities and to shed light on the overall probability density function of user contextual settings, tasks and user related moderating-variables. The overall goal of this activity is to provide quantitative evidence supporting the hypothesis of environmental/contextual consistency.

- Stage 5: Aims to identify a number of stable behavioural responses observed in consistent contexts/environments. The goal of this step is to prove that IR users respond consistently to reoccurring problems as long as they interact in consistent contexts/environments. Additionally at this stage, the nature of the observed consistency is investigated and the factors affecting it discussed in detail.

- Stage 6: During this stage the identified consistent behaviours are further shortlisted. The further investigation is limited only to the behaviours that can be taken advantage of during the personalised/context sensitive IR development. At this point the further evidence supporting the hypothesis of human rationality and consistency will be provided. Additionally at this stage the further analysis of factors affecting the shortlisted behaviours will be carried out along with the means of militating against their variability.

1.7 Contribution

The work presented in this thesis has four main outcomes contributing to knowledge:

- It identifies the conditions under which HIB is consistent.
- It proposes a new "Integrated Framework for HIB in-situ" that supports the process of analysis of Human Behavioural Consistency and allows for effective personalised/contextualised IR development.
- It applies the proposed framework to the group of work-related tasks of software developers validating the approach.
- It identifies a number of consistent behaviours that can be used in personalised/contextualised IR development targeting software developers.

The rest of this section provides the reader with an overview of the above outcomes and discusses how they contribute to the existing knowledge.

1.7.1 Nature of Human Information Behaviour Consistency

The major contribution of this research is the investigation of human information-related consistency. The work described in this thesis illustrates that humans are capable of making rational decisions regarding their information preferences and manifest them in a consistent manner. On the other hand the research also demonstrates

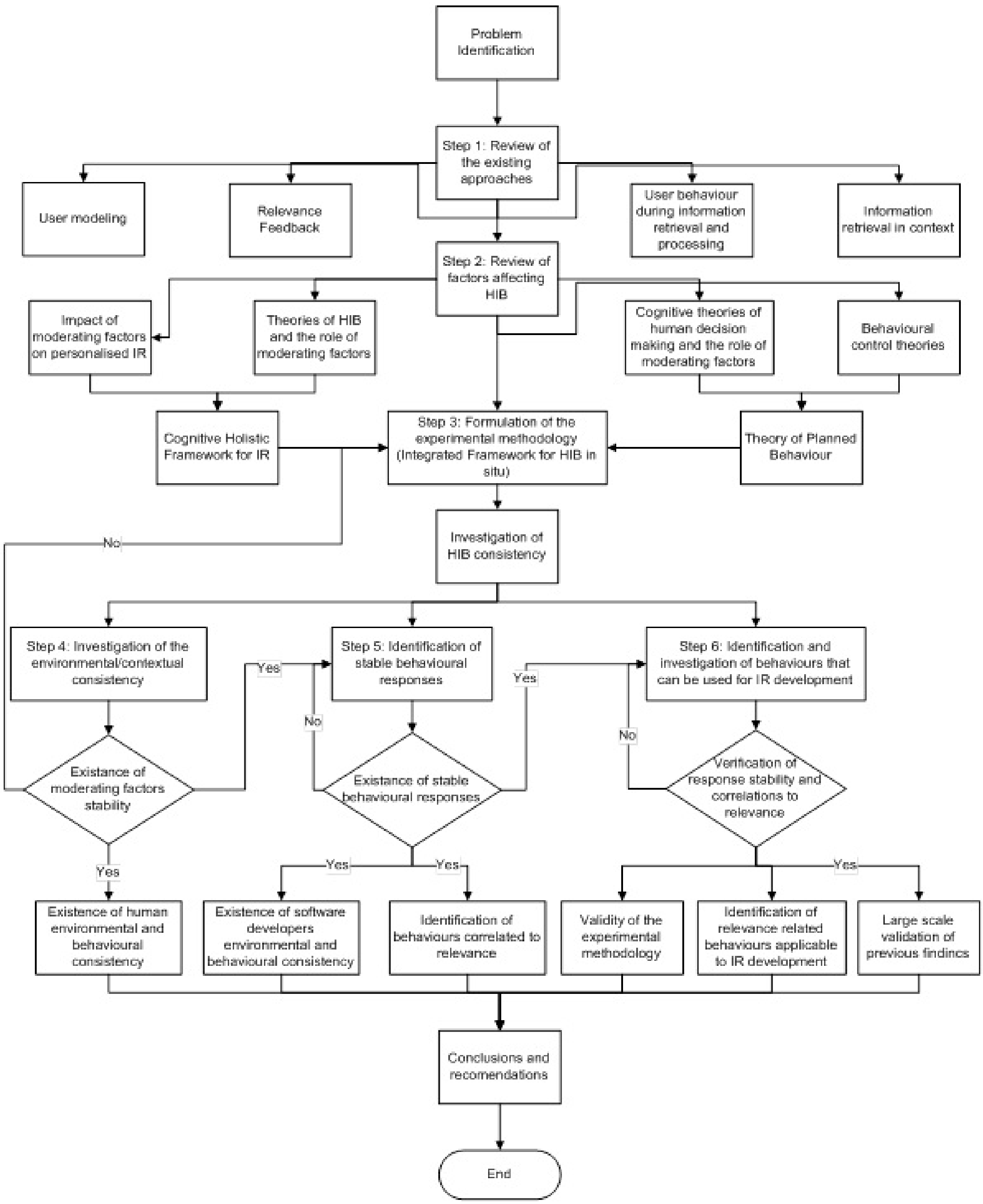


Figure 1.5: Research Methodology - The graphical illustration of the research methodology.

that due to the number of factors affecting the human decision-making process the observed behaviour varies significantly. In many cases the external observer can perceive this complexity as inconsistency. The thesis illustrates under which conditions HIB is visibly consistent and under which conditions high levels of behavioural variability can be expected.

The results of the user studies illustrate that the behavioural variability (HIB inconsistency) observed in the literature is caused by the changing perception of task, perception of the context and changes of the user. The results also indicate that there exists a number of tasks and contexts that are highly consistent in user life and to a high extent lead to consistent behavioural responses. The above observation is supported by the results of the user studies of this research that illustrate a non-uniform distribution of moderating-factors supporting the hypothesis of the context/environmental consistency. Similarly, for consistent tasks and context the results clearly illustrate non-uniform distribution of user micro and macro behaviours supporting the hypothesis of HIB consistency in those particular scenarios.

The research illustrates that this variability can be managed through the application of the proposed research framework, substantially increasing the levels of observed predictive validity. This makes not only the personalised/contextualised IR feasible, but also because of the simplicity of the proposed framework, financially viable. This is especially valid for commercial implementations focusing on tasks and user groups that are highly dependent on information.

1.7.2 "Integrated Framework for HIB in-situ" and its applicability to IR

One of the major contributions to knowledge is the proposed "Integrated Framework for HIB in-situ" and its application to case study of software developers. The proposed framework is based on two theories, namely: the Theory of Planned Behaviour and Holistic Cognitive Theory of IR. The integration of these theories allows the achievement of high predictive validity and simplifies the IR design process.

The proposed framework can be applied to address any HIB problem and support IR development regardless of the user group, task or context.

1.7.3 Consistency of software developers' Information Behaviour

The application of the proposed framework on the group of software developers illustrates consistency in developers' information behaviour. The results illustrate that software developers engage in finite and quite limited information activities which lead to a limited amount of information problems. The activities related to software development are not only very frequent but lead to a small number of information problems and information requirements. The results show that the way the problems are handled is very clearly defined and that the response of the developer to relevant information depends mainly on the problem classification. The results presented clearly illustrate the applicability of the proposed methodology to this particular user group illustrating its validity and ease of application.

1.7.4 Identification of strong implicit-feedback indicators for the population of software developers

The investigation of software developers' consistency led to the identification of behaviours that have high correlation to relevance and therefore can be used for implicit-feedback or collaborative-filtering. The shortlisted behaviours include:

- Direct code reuse practices at block level (distribution of cognition)
- Direct code reuse at subsystem and system level (opportunistic programming)
- Direct reuse of system level components e.g. libraries, classes, frameworks and bigger components (component reuse)
- Indirect code and information reuse practices at different complexity levels (block, subsystem, system)

The research investigated the user tendency to engage in the shortlisted behaviours, their correlation to information relevance and reviewed the moderating-factors affecting the variance. Similarly the research reviewed user preferences towards information objects used during the interaction as well as preferences towards user interfaces.

1.8 Thesis layout

The remainder of this thesis is divided into seven chapters each concluding with a brief summary together with the key ideas and concepts for ease of reference:

1.8.1 Chapter 2: Background

Chapter 2 focuses on the overall discussion regarding consistency of human behaviour. It discusses the key empirical research focusing on the problem of human consistency and rationality both in IS&R and Cognitive Psychology research. The goal of the discussion is to illustrate that a big proportion of our daily behaviours are consistent. If appropriate methods of behavioural aggregation are applied, then the overall behavioural trend is consistent and reasonable which is in opposition to the individual instances of behaviour. The chapter reviews the research devoted to the moderating-factor approach in predicting human behaviour and HIB and lists the advantages and limitations. It focuses on the analysis of the most frequently reviewed factors and compares the selected research from IR and Cognitive Psychology in order to provide the reader with the overview of the factors, their impact on HIB and the known correlations between the factors. The chapter also discusses the key models of HIB and investigates their components or moderating-factors and consistency with the empirical research.

1.8.2 Chapter 3: General Black Box Model

The goal of this chapter is to generalise the discussion about the HIB by introducing the reader with an abstract geometrical model. The model is introduced exclusively, in order to support discussion about the user consistency by the use of the mathematical vocabulary. This vocabulary is used not only in order to discuss the implications of HIB inconsistency but also to discuss the limitations of the moderating-factor approach to the prediction of HIB. Finally it is also used in the next chapters to discuss the features of the proposed "Integrated Framework for HIB in-situ".

1.8.3 Chapter 4: Achieving personalisation and contextualisation through behavioural consistency

Chapter 4 focuses on identifying the means to overcome the limitations of the moderating-factors approach. It discussed the two most representative (in the context of this work) research frameworks, examined from two different research domains: Cognitive Psychology and IR. The Ajzen's Theory of Planned Behaviour is compared to Ingwersen's Holistic Cognitive Framework of IR. Their strengths and weaknesses are discussed in the context of HIB consistency. This analysis focuses mainly on the capability of the selected frameworks to predict HIB and their applicability in the IR and IS&R research and development. Finally chapter 4 proposes a new framework for the research and development of IR and IS&R solutions. The framework is based on the two above mentioned methodologies the strengths and weaknesses of which are discussed in detail in the chapter.

1.8.4 Chapter 5: Software developers' behavioural consistency

The goal of chapter 5 is to illustrate how the proposed methodology can be used to analyse consistent aspects of HIB. It shows the way the methodology helps in the design and implementation of IR user interfaces, algorithms and personalisation features. The chapter focuses on a group of software developers and their task of code delivery. It reviews the rationale behind the selection of this particular user group and provides an overview of the theoretical work discussing the consistency of code related behaviour. It discusses the way the methodology is used to design user studies, that help to reveal the consistency and rationality of the developers' work-related efforts and prepares the reader for the analysis of the experimental results in the next chapter.

1.8.5 Chapter 6: Results

Chapter 6 discusses the empirical research analysing software related information needs and behavioural patterns. It provides the reader with the detailed description of the three user studies carried out on the group of software developers and a general information searcher population. The chapter summarises the key findings of the user studies and highlights the key findings of the proposed methodological steps.

1.8.6 Chapter 7: Conclusions

Chapter 7 of the thesis summarises the work carried out in this project and highlights the contributions to knowledge, strengths and limitations of that work as well as work planned in the immediate and distant future.

Chapter 1: Summary

Availability of effective Information Retrieval (IR) tools is essential for the success of modern information workers and as a consequence information focused business. The costs of non-effective IR to the global economy is significant and will rise with the availability of information.
Development of new personalised, context sensitive and task-focused IR solutions is one of the approaches aiming to address the above problem.
Modern contextualised and personalised IR utilities are very limited, especially in terms of the extent to which they are able to predict temporal relevance.
The majority of contextual and personalised IR systems are built on the assumption of human rationality, causality and consistency. The notion of human consistency is not well understood in IS&R as well as the HIB research.
Empirical research shows that human rationality and causality is rarely observed. Individual instances of human behaviour rarely have statistically significant correlation.
The development of business ready IR solutions requires understanding of the extent of human consistency and the factors affecting it.
This thesis investigates the above problem and proposes a research framework aiming to systematically support the development of personalised/contextualised IR solutions.
The thesis illustrates the application of the proposed framework on a group of professional software developers.

Table 1.1: Summary of chapter 1

2

Background

2.1 Introduction

As highlighted in Chapter 1 the notion of HIB consistency is a key for successful personalisation/contextualisation of IR systems. In order to create such systems:

- Human Information Behaviour has to be Consistent: It is not possible to model the relation between the past behaviour and future information needs of the user if this relationship does not exist.

- The nature of consistency has to be known: The relationship between user past behaviour and future information preference has to be formally described (e.g. as a function, heuristic or a formal model). If this relationship is significantly affected by external factors (such as context, user personality, etc.) they also have to be incorporated into this formal description or at least the magnitude of their impact on HIB prediction has to be understood. This is essential as not including the factor into the model introduces an error in behaviour prediction.

This chapter examines the existing literature in order to review the level of HIB consistency and investigates the factors affecting it. This chapter is composed of two sections that focus on the two key aspects of the problem highlighted above:

- Existence of Behavioural Consistency: This section discusses the history of the investigation of Human Behavioural Consistency. It illustrates how the consistency of human interaction with information was investigated throughout previous decades and how the work carried out by Lesk and Salton (1968) marginalised the problem of human unpredictability for years to come. It also clearly illustrates that the HIB is significantly more inconsistent than acknowledged in the research focused on personalisation/contextualisation of IR and that there is a need for further research aiming to describe the nature of human behavioural variability.

The discussion is followed with the review of the literature focusing on the problem of human consistency in everyday life. The psychological literature confirms that even the simplest behaviours are driven by a very complex cognitive process and as a consequence their repeatability is very limited. The review of the psychological literature confirms the findings of HIB research and clearly illustrates that individual instances of human behaviour are not consistent (Ajzen, 2005).

- Factors affecting Behavioural Consistency: Very frequently, the lack of observed behavioural consistency is explained by the existence of either external or internal factors moderating human behaviour. Factors such as context of IR interaction, task, user expertise or personality are believed to change the user behaviour to such significant extent that the actual observed IR interaction seems unpredictable. This section reviews the key literature both in HIB and cognitive psychology focusing on the investigation of moderating-factors and their impact on behaviour. It identifies an overwhelming number of factors that are proven to have statistically significant correlation with HIB and even greater number of interactions between those variables.

Overall the literature review carried out in this chapter illustrates that hu- man beings have a need to act consistently and live in a consistent environment as illustrated by balance theory (Heider, 1944); need for coherence (Kelly, 1992) and the need for logical consistency (McGuire, 1960). Even though this is the case, owing to the significant complexity of their environment and their own cognitive processes, their behaviour is significantly moderated by hundreds/thousands of variables. Without detailed understanding of all of those variables, it is not possible to predict human behaviour effectively. The empirical research both in cognitive psychology as well in general HIB confirms that, through very low predictive validity of its results.

This chapter illustrates that the extent of human consistency cannot be safely treated as an axiom the investigation has also to be a part of the IR research. The next chapter continues the discussion about the moderating-factors and their role in

HIB prediction. It introduces a mathematical model (General Black Box) which is used in the discussion of the feasibility of the moderating-factor approach and which also highlights a different direction for consistency research.

2.2 History of Behavioural Consistency Research

This section reviews the extent to which HIB as well as the everyday human behaviour is consistent. This chapter illustrates that individual instances of human behaviour are much more inconsistent than the majority of contextualisation/personalisation research acknowledges. It clearly illustrates that further understanding of when and to what extent human beings are consistent is critical for successful IR implementations. This section is divided into two subsections:

- The first subsection discusses the observed behavioural inconsistency in HIB research and its impact on implementation of IR systems and HIB research design.
- The second subsection discusses the observed behavioural inconsistency in our everyday behaviour. The second subsection illustrates the extent and complexity of the human inconsistency problem and the related research aiming to tackle the inconsistency challenge.

2.2.1 Consistency of Human Information Behaviour

One of the first (Ellis, 1996; Saracevic, 2007) and most frequently cited studies of information-related behavioural inconsistency was published in 1956 by Gull (1956). The goal of the experiment was not to assess the extent of user consistency or inconsistency between the users but instead to compare two competing indexing systems. What made the study unique was the fact that, in order to create the relevance benchmark for evaluation it engaged two independent groups rather than one. Each group was provided with a set of 98 queries/problems/requests and used the same set of 15.000 documents. Each group assessed the relevance of retrieved documents separately. The problem arose when the results of this manual assessment were collected and analysed. As commended by Cleverdon et al. (1966):

> " neither group was willing to accept the relevance assessments of the other group: rumour has it that at the end of the second day of discussion,

> the two groups were still arguing about the meaning of the first search question. No real blame can be fixed on those who organized the test because in 1952 it was not unreasonable to think that two groups of intelligent people would, without serious difficulty, be able to come to an amicable agreement as to which documents were relevant to a particular question."

The differences between relevance judgements were not trivial either. The first group selected 2200 relevant documents out of the indexed 15000 whereas the second group found only 1998 documents relevant. Surprisingly the overlap between the two selected sets was very low. The first group marked 1640 documents as relevant that the second group did not. Similarly 980 documents were marked as relevant by the second group and not by first. The further work by Claverdon resulted in some more compromises regarding the way relevance was assessed increasing the overall overlap but still the final agreement between the two groups was low and reached only 30.9%.

A significant amount of research followed the initial observations made by Gull. A group of studies was carried out by Resnick (1961); Resnick and Savage (1960, 1964) in order to investigates the human ability to consistently make relevance judgements and the impact of the information representation (the study used entire documents, citations, abstracts and index terms representing the documents) on the consistency of information selection (Resnick and Savage, 1964). The "intra-group" agreement varied between 54% and 70% and averages at 62%. A number of studies followed in the late 60's owing to the funding from the National Science Foundation. One of such studies was carried out by Rees and Schultz (1967). The goal of this study was to identify factors affecting human relevance assessment as well as the extent of inter user variability. The study involved 153 users from the medical domain divided into seven groups depending on their exact occupation or task group. Similarly to the previous study the "intra-group" agreement varied between 40% and 58%.

A similar study was carried out by Cuadra and Katter (1967); Cuadra et al. (1967). The study involved 230 students of psychology with varying levels of both psychological and search experience. The participants were split into four groups depending on their experience and asked to rate nine psychological journal abstracts against several short information statements. The results of inter user variability observed revealed an inter-user agreement (calculated as a correlation coefficient) varying from r = 0.41 to r=0.49.

Another important research focusing on the manual index creation process and the related consistency was carried out by Zunde and Dexter (1969). The study was carried out based on two groups of eight users carrying out the indexing process of 29 selected documents. The two groups varied with the level of indexing experience, eight of the participants being indexing experts and the other eight scientists related to the project. The results presented in Figure 2.1 illustrate that the consistency of the participants varies depending on the measure selected and ranges between 0.24 and 0.41.

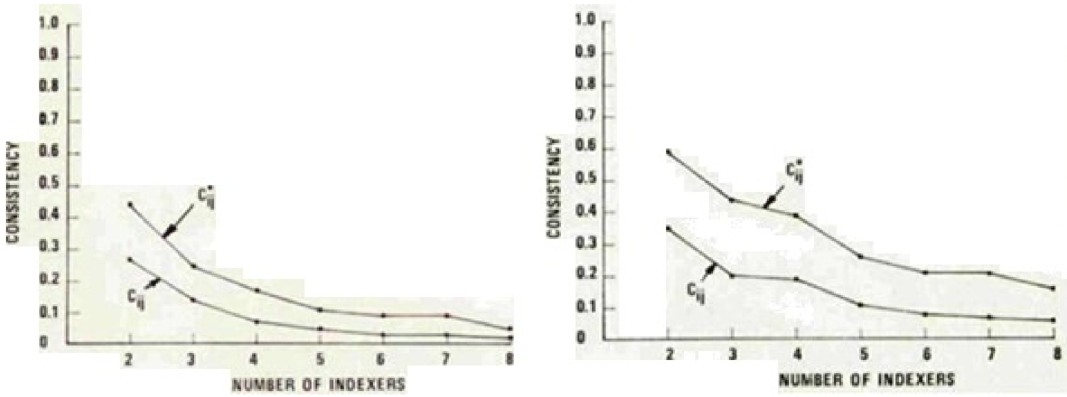


Figure 2.1: Level of user inconsistency (Zunde and Dexter, 1969) - The decrease of inter user consistency in index preparation depending on number of users (Two index consistency measures and two averaging techniques).

The further analysis also suggests that the consistency is inversely proportional to the number of the indexers in the selected group. The results of the studies discussed above did not change the way human behavioural consistency was perceived and treated by the IS&R research. Even though the above examples provide clear evidence for HIB inconsistency the evaluation of relevance in laboratory experiments was more and more frequently carried out based on a single evaluators group and the design leading to observation of inconsistencies was avoided.

The study that sealed the fate of human consistency research was carried out by Lesk and Salton (1968). The experiment hypothesises that the variation of user relevance assessment does not have an impact on the relevance algorithm evaluation process. The experiment involved a group of 8 users from the library research field assessing the quality of the results of 48 queries based on a collection of 1268 abstracts stored in the developed system. During the first iteration the group was asked to author 48 search queries (each participant constructed 6 queries) and assess the quality of the retrieved content from the information problem author's perspective. The relevance was assessed dichotomously. During the second iteration the same group of users were asked to assess the relevance of the 6 queries that they did not author. The results reveal only 30% of agreement between the judges and significant individual differences between the judges regardless of their very similar background.

Even though the agreement between the users observed was very low, Lesk and Salton (1968) have hypothesised that this will not have an impact on the efficiency of the IR system evaluation process. During this study they have formulated two hypotheses that have changed the way human consistency was reasserted for the next decades. The strong hypothesis states that the variations of HIB during the laboratory IR system evaluation will not affect the assessment of system performance. The weak hypothesis states that the same variability will not affect the assessment of different retrieval methods.

The research carried out by Burgin (1992) was designed to verify Lesk and Salton's strong and weak hypothesis. The experiment involved the use of 1239 biomedical documents and their evaluation against 100 queries based on 6 different document representations. The experiment involved four sets of users:

- a very experienced domain expert and an author of the queries,
- nine paediatric faculty practitioners,
- four post-doctoral fellows and
- a medical bibliographer with significant search experience.

The results show the level of agreement varying between 39.9% and 55%.

The experiment carried out by Shaw et al. (1991) had the same methodology as the Burgin's study reported above. It took advantage of the same database and the same query collection, grouping the users in the same way as well, but took advantage of a different indexing technique. The overall agreement rate between the four groups engaged in the study was reported to be equal to 40%.

The above mentioned research started shaping a gap between the user and system-oriented research which is still very apparent until today. System-centred research continued the work using the laboratory approach (based on the Cranfield, and what would later become TREC methodology) aiming at automatic evaluation of IR systems and algorithms. The laboratory focused work ignored the notion of user behavioural inconsistency completely, resulting in the user being removed from the research design and substituted with various test corpuses. The user-focused research, on the other hand, concentrated on the HIB and factors affecting it, to some extent losing track of real life system implementations and development methodologies. The paths of those two research groups diverged, with the system focused studies continuing to deliver more effective IR solutions (it is said that the IR system efficiency doubled within the first 6 years of TREC). The user-centred camp then focussed on the complexities of user behaviour, including the nuances of human consistency.

One of the most representative user-centred studies on the factors affecting information behaviour was carried out by Saracevic et al. (1988). The goal of the study was to capture and classify the factors significantly impacting human IS&R with the main emphasis on cognitive context and human decisions and interactions involved (Saracevic et al., 1988). The experiment involved 40 participants (19 faculty, 15 graduate students, and 6 professionals from industry) taking part in multiple stages of experiment of various natures (interview, questionnaire, search attempt and, a relevance estimation exercise) and 5 professional searchers carrying out the retrieval tasks. The study investigated the problem of user search term selection and item selection consistency. The average inter-judge agreement reported averaged at 27% and more than 55% of responses having an agreement of less than 25% (Saracevic et al., 1988). The consistency of document selection was reported to be even lower. The average agreement for selection of relevant and partially relevant items was equal to 18% and almost 60% of all items selected only had agreement at levels from 0% to 5%.

The experiment carried out by Saracevic and Kantor also attempted to find the correlation between the selection of the search terms (i.e. attitudes towards the problem expressed through a query) and the information selected. The results not only did not manage to show statistically significant correlation but illustrated that the data is capable of explaining only 2.5% of variance. A significant number of other studies followed aiming to illustrate the significant extent of user behavioural variance and limitations of laboratory / TREC approaches. The experiment carried out by Lee et al. (2006) used 10 users with at least 4 years of search expertise to assess the value of the search results. Each participant was presented with a number of search result lists, two at a time and was asked to assess their pairwise relevance to the topic at hand.

The precision value of the result list varied, which was established based on previous relevance judgement. In each of the iterations, the users were presented with one result list of 50% precision and a second one where precision varied between 30% and 70%. The recognition of correct lists reported varied between 14.6% and 31.2%. The overall inter-judge agreement was reported to be equal to 24%.

A number of user information-related consistency studies was carried out by Sormunen (2002); Vakkari and Sormunen (2004). One of the goals of the study was the identification of relevance assessment variation inconsistencies between the relevance assessments captured when creating the TREC test data and those captured during the experimentation. The study carried out by Sormunen (2002) focused on TREC-7 containing 8 collections and involving 9 master students reassessing the 5271 documents from those collections against the selected topics. The results demonstrate very significant differences in the assessment process. 25% of the documents that were assessed as relevant in the TREC collection were assessed as highly relevant by the research participants whereas 36% of TREC relevant documents were given marginal relevancy. Discrepancies were also identified among the documents that in the TREC data collection were marked as irrelevant, with 1% of those marked as relevant or highly relevant.

A similar study to the procedure adopted from Sormunen (2002) was carried out by Vakkari and Sormunen (2004) engaging 26 students to re-evaluate the TREC-9 data. The results show that only 45% of documents marked as relevant in the TREC collection were assessed as highly relevant by the experiment participants.

Yet another example of a study investigating the consistency variations by referring to the TREC data was carried out by Voorhees (Voorhees, 1998, 2000, 2001; Voorhees et al., 2005). The experiments focused on the impact of inconsistency in judgement on the laboratory IR research. The variation of user consistency measured as an inter-judge agreement varies between the studies from 30%-33% (Voorhees, 1998, 2000) to 34% (Voorhees, 2001).

The studies reported above are just the most representative research examples of the domain, and a significant amount of other research that focused on relevance assessment consistency, indexing consistency and information object selection consistency are available. This brief review illustrates that the reported levels of inconsistency varies significantly ranging from as little as 3.5% (Haynes et al., 1990) to as much as

60%-70% (Resnick and Savage, 1964). Overall though, the overlap or behavioural consistency seems to be observed at a level of 30% which is a value curiously similar to the .30 threshold reported by similar research in the cognitive psychology domain (Wicker, 1969).

2.2.2 Consistency of General Human Behaviour

It is very important to notice that the problem of human behaviour inconsistency is not unique to IS&R research. The ability to predict human behaviour is much desired throughout multiple research disciplines therefore a very mature research focusing on the issues of consistency exists. The psychological findings about human behaviour consistency especially are very relevant in this context as HIB is just a subset of human everyday behaviour and the same cognitive mechanisms are used when solving IR related problems. It is therefore very valuable to review the findings of psychological literature and review how it dealt with apparent inconsistencies, especially in the context of its 100 year long tradition.

As expected the psychological research found very little consistency between the individual instances of human behaviour supporting the findings identified in the IR domain (refer to the previous section). The lack of the observed consistency is especially significant if the behaviours are observed in different contexts or settings. Importantly the observed levels of inconsistency were reported to be very high even for behaviours significantly simpler than the average IR behaviour illustrating the complexity of the problem.

The psychological literature is not only very valuable for the purpose of investigation of HIB consistency but also directly applies to the problem of contextualised and personalised IR development. The investigation of the relationship between the attitudes declared by the user study participants and their observed behaviour directly impacts the way we perceive systems based on explicit sources of relevance-feedback. Similarly the research focusing on the existence of correlation between individual instances of behaviour has direct implications for the design of systems based on implicit-feedback analysis and creation of user profiles.

One of the first and most often quoted studies discussing the extent of consistency of human responses was carried out by LaPiere (1934). The experiment investigated the extent of racial discrimination experienced in various interaction settings. In early 1930

LaPiere took a Chinese couple to 251 restaurants, hotels and other similar locations in order to establish the extent to which the couple will experience racial discrimination manifested through denial of service. Out of 251 establishments only 1 denied the service. Following this, the subgroup of 128 establishments was contacted by letter asking whether they would accept a guest of "Chinese race". What is interesting in the study is the response to the written enquiry which in 90% of cases was negative. This drastic inconsistency which seems to be triggered exclusively by a trivial contextual factor (i.e. the presence of a discriminated person or group) inspired an entire family of behavioural consistency research.

A very similar study on racial discrimination and selection of discriminating behaviour was carried out by Minard (1952) over four weeks during the summer of 1946. The study consisted of a number of interviews with the miners as well as observation of their interaction in work and off-work environments. The results illustrate significant inconsistencies in the white miner racial behaviour. Minard observed a very high level of integration with black co-workers in the work environment but very widespread segregation in the community. The behavioural inconsistency was observed across 60% of miners and only 40% either did not racially discriminate miners at all, or discriminated them consistently regardless of the settings.

The study carried out by Mischel and Peake (Mischel and Peake, 1982*a*,*b*, 1983) investigates three different types of behavioural consistency:

- Temporal Stability: referring to the consistency of the same behaviours in the same settings/ contexts.
- Cross-Situational, Uni-Modal Consistency: referring to the consistency of the same behaviours in varying settings/contexts.
- Cross-Situational, Cross-Modal Consistency: referring to the high level consistency across multiple different behaviours from a same class/group across different settings/contexts.

The experiment involved observation of 19 different behavioural tendencies observed in repeating and different occasions. The tendencies included a variety of behaviours such as thoroughness of note taking during classes or personal appearance. The results illustrated very low level of consistency for all of the types of consistency. The correlation coefficient for Temporal Stability was reported to be merely .29 and for the Cross-Situational, Uni-Modal Consistency .08. The consistency measures based on the behavioural aggregates were reported to be significantly higher for Temporal Stability reaching .68, but a significant increase was not observed for other types of consistency. The average correlation among all types of behaviours was .13.

The above inconsistency of behaviour is not only observed between various instances of the same behaviour but also to a significant extent between the behaviour and our attitudes towards those behaviours expressed verbally. This type of observed inconsistency is especially important in the context of IR laboratory experimentation design as it identifies a difference between the user's verbal attitude towards information and his further assessments through its use.

A very frequently quoted example of research investigating the correlation between user general attitudes and behaviour is the problem of racial prejudice and discriminating behaviour (Ajzen and Fishbein, 1977; Duckitt, 1992, 1994; Kraus, 1995; Wicker, 1969). The original review of the verbal and overt racial responses carried out initially by Wicker (1969), further expanded by Ajzen and Fishbein (1977) and Duckitt (1994) was used for years as a basis for the discussion on attitude to behaviour relationship. The discussion in all of the mentioned cases leads to similar conclusions. It seems that the relationship between the verbal indicators of prejudice and behaviour bears very low correlation. The review of literature carried out by Ajzen and Fishbein (1977) concludes that out of 54 studies reviewed that aimed to predict behaviour on the basis of attitudes 25 obtained non-significant results and the rest of them rarely showed a correlation of 0.4 and above. The original review carried out by Wicker (1969) reported the average correlations not higher than 0.3.

One of the most recent and most comprehensive meta-analysis of the literature devoted to the prediction of human behaviour based on declared attitudes was carried out by Kraus (1995). The results of his analysis (presented in Table 2.1) reveal that the Pearson's p of a correlation has a normal distribution with a mean of 0.38 and median of 0.33. The results illustrate also that the attitudes explain 14% of the behavioural variance.

The above review, in the context of IR, suggests that explicit measures of information relevance, especially in the controlled laboratory environment, can be carrying a very significant error. This error will be even bigger especially if the contextual/social

Statistic	*Effect Size (ABC)*	*Significance Level (Z)*
Maximum	.91	38.37
75th percentile	.50	5.57
50th percentile (median)	.33	3.57
25th percentile	.22	1.99
Minimum	–.10	–0.76
Mean	.38[a]	4.62
Standard deviation	.21	4.96
Kurtosis	–.07	23.38
Skewness	.21	4.05

NOTE: ABC = attitude-behavior correlation.
a. Calculated by converting each *r* to a Fisher's *Z*, averaging the Fisher's *Z*s, and converting the average *Z* back to *r* (Rosenthal, 1991). All other results in this table are based directly on the Pearson's *r*s.

Table 2.1: The meta analysis of attitude-behaviour correlation research (Kraus, 1995)

settings of the experiment differ significantly with the real life IR situation. The results reported above confirm that human beings seem to expose significant behavioural inconsistencies whenever their context of interaction changes.

Another very important question in cognitive psychology relates to the behavioural consistency of people sharing the same personality traits. This problem is equivalent to the problem of user profiling in IS&R. There exists a significant amount of research investigating the correlation between various user traits such as personality factors and the observed behaviour. The key research focuses on the analysis of the big five personality factors and observable behaviour.

One of the most representative studies discussing the personality related consistency was carried out by Paunonen (2003). The study involved two groups of students (n=276 and n=142) who were requested to fill out a personality questionnaire and a behavioural report form. The behavioural form measured the user engagement in all of the 27 behaviours investigated in the study. The results illustrate that only a marginal number of generic behaviours were successfully correlated with personality measures for both groups. The correlation coefficients were reported to be of low value rarely above 0.3. A significant proportion of the behaviours were not successfully correlated with personality for both groups.

Another very representative study on the correlation of personality to observed behaviour was carried out by Markey et al. (2004). The study was based on the observation of the everyday behaviour of 94 ten year old children. The study involved the observation of 63 types of behaviour which were further encoded and correlated with the results of the personality tests. The analysis of the results reveals that even though a significant proportion of the correlation scores were statistically significant only in two cases ("Says negative things about self" versus "Neuroticism" and "Interviews his or her parent(s)" versus "Extravertism") was the correlation higher or equal to .3 or lower or equal to -.3.

The correlation values reported in this research are very characteristic and reoccur frequently across multiple experiments devoted to user consistency research. The observed re-occurrence of correlation values between .20 and .30 was reported frequently in the literature (Mischel, 1969; Wicker, 1969). In a book entitled Personality and Assessment published in late 60s, Mischel (1969) reviewed much of the research on personality and its correlation to behaviour and coined a term of a "personality coefficient" relating to the observed trend. In his book Mischel observes that it is very rare to observe correlation of a magnitude above .30. This means that the attitudes and human personality traits account for less than 10% of variance by themselves. This observation holds true not only for the big-five personality factors but seems to reoccur for various personality and attitude classifications used. This on the other hand indicates how significant is the influence of the outside environment on our decision-making process. This also supports the magnitude of results obtained in the IS&R research discussed in the previous section.

Overall the empirical research illustrates very little evidence for the existence of behavioural consistency between single instances of user judgements and actions as well as between user traits and their future behaviour. As observed by Ajzen (Ajzen, 2005; Ajzen and Fishbein, 1977), by the end of 1960s the personality and social psychology research had noticed that the traits and attitudes alone are poor predictors of behaviours predicting marginal proportion of variance.

2.2.3 Conclusions

The empirical evidence presented in this section clearly illustrates that the consistency of human behaviour should not be taken for granted. On the contrary the results of both the psychological research as well as IS&R research suggests that individual instances of human behaviour are not correlated. Icek Azjen, when reviewing the issues of human consistency even noted that it seems that human consistency and repeatability is "more apparent than real" (Ajzen, 2005) stressing the fact that even though we like to perceive ourselves as rational and consistent the empirical evidence seems to prove otherwise.

There are significant theoretical foundations (Abelson, 1968) that support the human need for consistency in everyday life, including balance theory (Heider, 1944), need for coherence (Kelly, 1992), and need for logical consistency (McGuire, 1960). However the research indicates that the need for consistency is not a main driving force in human decision-making process but rather a mere preference (Zajonc, 1968). No anxiety is present when we make inconsistent decisions and neither is there a mechanism that would create a compelling need to strive towards consistency. As a consequence and in contrast to the common intuition the empirical research has failed to demonstrate human consistency and the predictive potential of human characteristics and past behaviour on a single occasion.

The literature review carried out by Mischel (1969) and Wicker (1969) illustrate a very weak correlation of attitudes to human behaviour rarely above .30 and often very close to zero (Mischel, 1968). A review carried out by Ajzen and Fishbein (1977) covered 109 distinct investigations looking at the relation between user global attitudes and individual instances of behaviour. Out of those studies only 25 obtained statistically significant results and from among those only few demonstrated correlation coefficient in excess of 0.4. Similarly the correlation of aspects of attitude to behaviour was very low usually ranging between .20 and .30. More interestingly this correlation, frequently referred to as the "personality coefficient", is consistent in the literature and is being observed in almost all (if not virtually all) research aiming to identify the correlation between any measurable personality traits and external responses (Mischel, 1968). The review of the above literature leads to the conclusion that only a very small proportion of behavioural variance can be explained by human attitude to a given action (i.e. selection of a relevant document) and user personality traits (Wicker, 1969). For example a review carried out by Hunt (1965) (p. 81) reveals that only 4% to 25% of behavioural variance can be explained by the personality traits. As a consequence the research in personalized IR, especially the one focused on explicit or implicit forms of feedback and analysis of past interactions (for example through the log analysis) should also observe a significant inconsistency in the human content selection process and relevance assessment. As this review illustrated this is indeed the case.

The results on user consistency in IS&R are virtually identical reporting inter and intra-user agreement of .20 to .30 for bigger groups and more complex variations of contextual settings demonstrating the same trend in this particular branch of social research. What is more interesting is the fact that the conclusions regarding human consistency and its extent are identical in both cognitive and social psychology as well as in the IS&R and HIB research.

2.3 Moderating-factors affecting Behavioural Consistency

The variations of human consistency in IR and HIB is often rationalised by the impact the user interaction "Context" or "Task" has on the decision-making process. A significant proportion of research community reached a conclusion that the same tasks / IR problems are performed in different situations, different social settings and on different individual predispositions. Different features of those situations and settings can affect the performance of a particular behaviour, search strategy or result assessment. Moreover they believe that certain personal traits can not only have a standalone effect on human behaviour but also interact with the situational factors moderating individuals' behaviour even further. For example the user domain expertise will not only affect user performance when carrying out tasks within the domain but will also affect users' perception of complexity and stress level. This leads to a change in self-motivation and the will to proceed with finding the solution to a problem. Considerable research in psychology illustrates similarly that even the slightest changes in the context of the problem can change the outcomes of the decision-making process (Tversky and Kahneman, 1992). As a consequence the same problem in the domain of psychology led to the same conclusions regarding contextualisation. The early 70s and 80s introduced notions of interactionism and contingent consistency or MODE (Fazio and Towles-Schwen, 1999) approaches. Those approaches are focused on the impact of situational circumstances (Monson et al., 1982), characteristics of the individual (Schneider and Chein, 2003; Snyder, 1974) and secondary characteristics of human disposition (Norman, 1975) to increase the level of correlation between user attitudes and individual behaviours. The factors affecting human information behaviour and moderating the correlation between the human attitudes and actual behaviour (moderating-variables) were intensively investigated during the last several decades. The results of this research illustrated that in specific situations the moderating approaches can increase the observed correlation coefficients. The moderating-variable approaches are therefore logically consistent with the modern IR and HIB attempts introducing context into the relevance prediction algorithms.

This section reviews the modern moderating-factor approaches both in HIB as well as in general psychological research. The goal of the review is to capture the key moderating-factors affecting the consistency in order to assess the challenges of the moderating-factor approach.

2.3.1 Moderating-factors: HIB research

There exists a body of literature discussing factors affecting HIB. The majority of research on behavioural moderation is captured through various models of Human Information Behaviour (Case, 2007; Fisher and Julien, 2005; Spink and Cole, 2005; Spink and Heinström, 2011). The available models describe the process from various perspectives focusing on the motives of IS&R and the chronological order or the overall shape of the mechanisms involved in the process. This section will focus exclusively on the key models, analysing the factors affecting human information-related decision-making process and enriching the discussion of the moderating-factors approach.

2.3.1.1 Sense-making methodology

"Sense-making" is one of the most widely recognised and applied models of HIB and a methodology for experiment design and implementation (Dervin et al., 2003). White and McCain (1998) consider Dervin's work as one of the most recognised in Human Information Behaviour research.

"Sense making" methodology strictly speaking is not a model of HIB but instead it comprises of a set of meta-theoretical/ philosophical assumptions about the world and human cognition of the world. As a consequence the 'Sense making' methodology is not intended to explain mechanisms of human cognition nor is it intended to directly guide the HIB experimentation process.

According to Dervin et al. (2005) the goal of the methodology is "to assist researchers and practitioners in formulating approaches that advance our understanding and help us avoid the creating of mythical users who do not exist, creating instead, as if seen through tiny keyholes, useful pictures of "real" users with whom information systems can more meaningfully intersect". The "Sense making" methodology strives to guide a process of cognitive theory formulation and tests (Fisher et al., 2003; Vakkari and Sormunen, 2004) and because of its meta-theoretical nature it is by definition untestable.

Dervin's work on "Sense making" is primarily founded on the observation/assumption that has to actively analyse the surrounding world all the time making sense of it. The methodology is based on the assumption that human perception of the world is discontinuous or full of gaps. According to the methodology human beings encounter those gaps all the time. There exists a continuous discrepancy between the objective world and what we know about it. This gap exists between:

- People - e.g. gaps between the intended message and the way it is interpreted
- Things - e.g. gaps between the objective properties of object and their perception
- Spaces - e.g. gaps between the perception of the world in different places and situations
- Times - e.g. gaps between how we perceive the world now, in the past and future

Whenever a human encounters such cognitive gap the sense making process is engaged intending to fill the gap with information. The essential observation of the sense making methodology is that this process is carried out in a specified point of time, in a clearly defined place and context by a certain human being or group of people. As a consequence the information in the 'Sense-making' methodology is always a subjective construct of a certain human being facing a gap in his cognition at a certain point of space and time.

The sense making methodology also makes a significant number of assumptions about the human being under investigation. Dervin Dervin (1999) believes that "the human is conceptualized as: centred and decentred; ordered and chaotic; cognitive, physical, spiritual, and emotional; and potentially differing in all these dimensions across time and across space". The methodology does not assume that the human being is static/consistent in time and space and acknowledges the ever-changing nature of human cognition.

Figure 2.2 provides a core Sense-making metaphor. It illustrates the individual's ongoing travel through the information space, encountering an information gap, and filling it through the sense-making process.

This image has been removed

Figure 2.2: Sense Making - Dervin's Sense-making methodology (Dervin et al., 2003)

In her Sense-making methodology Dervin describes a human being as an ultimately causal (mechanistic causation) creature whose actions/activities are effects of the information collection process. The sense making behaviour is a direct response to the human being encountering a gap in his cognitive description of the world. Similarly, the finding of the information leads inevitably to the bridging of the gap and allows the user to continue the travel through cognitive space. This central assumption of human ultimate causality is very appealing as it allows for the discipline to be treated as natural science and promises to make it possible to elucidate the mechanisms of HIB. The approach explains the complexities of the observation of this ultimate causality with the ever changing nature of the human and its environment affecting the nature of the gaps themselves.

This methodology used since early 1980s, is a specification of the generic "Problem Space Theory of Problem-Solving" (Newell and Simon, 1972). Similar to Newell's and Simon's work, Sense-making theory notices that in order to solve a given problem a Mental Model of the problem is being created. The problem is then transformed into a number of problem states composing together a problem space. Figure 2.3 illustrates a number of different problem spaces that represent different ways a human can perceive the problem. The more constrained the problem (i.e. the more defined is the task) then the clearer is the choice of the appropriate problem space. Ill-defined problems often lead to the situation where the incorrect problem space is chosen that sometimes leads to the situation where there exists no path leading to a solution.

The problem space is constructed of a number of states including: initial state and the goal state as well as a number of legal operations and constraints that can be used to transform the temporary states into a goal state. Humans use existing schemas for transition from one state to another (figure 2.3) and in the case of the lack of a schema a search process starts which creates a new different problem.

This image has been removed

Figure 2.3: Problem-Solving Representations - Problem-solving mental models (Robertson, 1999); Gick's schematic diagram of problem-solving (Gick, 1986)

Dervin's Sense-making methodology is therefore a specification of the "Problem Space Theory of Problem-Solving" for the problem of HIB and the clear definition of the key elements of the user mental model in the IS&R situation.

Such approach to the problem of HIB has been very frequently criticised. One of the predominant critics was made by Rosenbaum et al. (2003) in the 2003 ASIST proceedings. Their work, discusses during the ASIST panel called "The death of the user", reviews a number of both practical as well as fundamental theoretical issues related to such a user-centred approach. In their work they criticise the fundamental assumption of this approach, that is the assumption of human mechanical causality. They believe that in the world where human beings are forced frequently to act under uncertainty, and are endlessly bombarded by information gaps, there is no anxiety associated with having those information gaps unfilled, and as a consequence not all of the gaps lead to sense making/ information seeking activity (Davenport, 2010). The introduction of the concept of the 'gap' or as discussed by the "anomalous states" notion of Belkin (1980), "raises more questions than they answer" (Day, 2011). The criticism goes significantly beyond the above and notices the fact that human beings are not mechanically causal by nature, and do not (as suggested by Dervin's work), make sense of the surrounding world in isolation, but rather take part in a "pragmatic activity of social and cultural positioning" (Day, 2011).

The above raises a more fundamental (and partially philosophical) question of the role of the information gap, anomalous states of knowledge, information intentions or more widely speaking the role of the concept of plans/purposeful action as such. Are the plans that we create to solve our everyday problems prerequisites for the action to occur, or are they simply a consequence of humans understanding their action, or one of the key resources used to support the action (Schmidt et al., 1978; Suchman, 2007, 1987). Taking into account the properties of the environment in which human kind evolved which are full of uncertainty and where the decisions need to be made based on inaccurate knowledge, the ability to respond dynamically to the situation, and our changing perception of the world, seems to be critical for our survival. In our daily life human beings do not always have the luxury of being able to define their information gaps and interpret them within the context of their environment. Humans indeed formulate plans that support their behaviour but those plans are "necessarily vague, insofar as they must accommodate the unforeseeable contingencies of particular situations" (Suchman, 1987). Moreover, as empirical research suggests and as discussed earlier not all of the plans that are formulated are also transformed into action. Similarly not all of the actions are preceded by any explicit plans and some plans lead to a lack of action (Allen, 1984). Suchman (2007) suggests that plans can potentially be "constituent as an artifact of our reasoning about action, not as the generative mechanism of action", (which would also be very consistent with "Self-perception" theories (Bem, 1967; Condry, 1977)). The above make the plan, information gap, anomalous state of knowledge and other similar concepts a very complex unit for further analysis of HIB.

The work is also criticised for its inability to effectively support the development of commercial IS&R and information management systems. As discussed during the 2003 ASIST conference: "the cash value of the concept of the user has reached a point of marginal or diminishing returns" Rosenbaum et al. (2003). It is widely believed that user studies based on the paradigm of the user as defined by Dervin (but as it will illustrated also Belkin, Wilson and other) do not scale well to wider user groups such as organisations or industries. As a consequence systematic design based on the proposed methodologies is claimed to be very challenging but also the actual systems designed on its basis are claimed to have limited usefulness Gutwin and Greenberg (1998); Guzdial et al. (2002). Rather than focusing on the key and recurring actions of systems design, they focus on the mental models of individual users and examination of factors affecting their decision-making process.

The above leads to the identification of thousands of factors which occasionally affect human behaviour and do not help in the identification of key trends in HIB that can be automated or supported by computer systems.

2.3.1.2 Belkin's Anomalous State of Knowledge

The approach of Belkin (1980), is another example of Cognitive Approach to LIS and HIB. It is based on the same fundamental assumptions about the human cognition as the previously discussed model proposed by Dervin (1983), but also shares a lot of commonality with other research focusing on the concept of "Information need" (Dervin et al., 2003; Taylor, 1962; Wersig, 1979).

This image has been removed

Figure 2.4: Belkin's ASK - Belkin's Anomalous State of Knowledge (Belkin, 1980)

Belkin's work (as illustrated in figure 2.4) acknowledges the fact that information seeking behaviour is triggered by insufficient information about the surrounding world which he refers to as an "Anomalous State of Knowledge" or ASK. What is more important his work acknowledges and heavily stresses the fact that the direct consequence of being in the state of ASK is also a limited ability to describe the information need. This problem of "non-specifiability of need" (Belkin, 1980) can has two dimensions:

- cognitive - founded on human incapability to conceptualize the path that will lead to the satisfaction of the need
- linguistic - based on inability to formulate in machine-recognisable language for example due to the limited vocabulary

As a consequence the proposed model suggests the development of systems that are based on "a statement of what the user does not know" (Belkin et al., 1982), rather than a statement describing the problem. Belkin believes that the future IR systems should be highly iterative engaging the human user in a dialogue that will not only help him to reformulate the need (so address the linguistic challenge) but also provide systematically information to allow the user to refine the mental model of the problem (addressing the cognitive challenge).

The above observation contributed significantly to the design of modern computer systems. Modern IR systems as well as the general structure of the Internet attempt to compensate for the users' inability to formulate the search problem through the employment of techniques such as Latent Semantic Indexing (compensating user's linguistics and cognitive capabilities), but also through providing a dense link structure enabling information foraging (in order to support reformulation of the mental model) or multi-faceted representation of information.

The criticism of Belkin's work is not focused on the above mentioned observation of human inability to describe their problem, but primarily on the underlying assumption of ASK being the primary trigger for the Information Seeking Behaviour. Modern research correctly identifies the fact that very frequently humans engage in seeking activity even though they do not have any specific information need (e.g. because of boredom). Similarly they notice that not all ASK's lead to an IR response. Humankind have evolved to efficiently cope with uncertainty (caused by almost infinite complexity of the surrounding world) and effectively manage the associated risk. As a consequence majority of even most critical decisions are made under uncertainty and do not lead to Information Seeking or even anxiety.

Another major criticism lies in the fact that the proposed meta-theory does not significantly simplify the design of HIB experimentation nor makes the development of IR utilities any simpler or more systematic. It is more of a philosophical assumption about the nature of human cognition which is difficult to translate through any particular research or engineering effort.

2.3.1.3 Everyday Life Information Seeking (ELIS)

The model proposed by Savolainen (figure 2.5) is one of the first models to stress the importance of social and cultural factors that affect the human information-related decision-making process (Savolainen, 1995). The model was aimed to complement the very popular research in work-related information seeking, hence its name: "Everyday Life Information Seeking". The proposed model is based on the concept of the "Way of life" (heavily inspired by Bourdieu's Habitus) which can be described as a general way of thinking about our everyday concepts, shaped by our cultural and social background. The "Way of life" in this model is managed by the "Mastery of life" element which is composed of several separate mechanisms:

- the user personality trait (i.e. "Main type of mastery of life"),
- the cognitive ability to solve problems and its attributes (i.e. "Problem-solving behaviour"),
- the ability to interpret the situation at hand accordingly.

The section on the right of the diagram represents the "Capital of Life" which is held by the human being or the environment and which affects both the "Way of Life" and "Mastery of Life Function".

The idea of the model is to provide a holistic framework for the description of social and psychological factors affecting the user everyday information-related decision-making process. As such it is very abstract in terms of the listed factors but also provides a very flexible framework for further research and illustrates the richness of the process.

This image has been removed

Figure 2.5: ELIS model - ELIS model of HIB (Savolainen, 1995)

Savolainen's Savolainen (1995) model is heavily inspired by Dervin's work. It primary objective is to translate the concepts of the information gap into the domain of everyday information interaction, which is much less defined and constrained than the work-related information activities. As such, it focuses on the main abstract factors bounding and controlling the seeking behaviour as well as their interaction. Similarly to Dervin's Sense Making (Dervin et al., 2003), Savolainen's work is also heavily user-centred making the user a centre of the information system analysis.

From the perspective of this research and the discussion of moderating-factors affecting HIB, ELIS model introduces yet another set of factors significantly affecting the information seeking process. Similarly to Dervin's or Belkin's work the proposed model is a meta-theory guiding the process of further theory design. It is not easily translatable into the research or system design process.

2.3.1.4 Wilson's model

Wilson (1999) model is very important for the purpose of this research as it explicitly focuses on a number of moderating-variables in this instance referred to as "Intervening variables" (figure 2.6). It recognises the need to analyse user behaviour, related to the information need in a perceived user context. From that perspective and also because of its overall structure it is very similar to the "Perceptual Control Theory".

This image has been removed

Figure 2.6: Wilson's models - Third Wilson's model of HIB (Wilson, 1999); Generalised Wilson's conclusions of the HIB field analysis (Wilson, 1999)

The model not only lists the core processes involved in HIB but also groups them in a way that focuses on the information. The model treats the user as an information control system which responds to the external environment, modifies its internal state by learning and then responds modifying the external information environment. Such a shape of model is not coincidental, which is underlined by Wilson when he states that information-seeking behaviour should be perceived as a goal-determined and iterative process (Fisher and Julien, 2005)).

The model evolved through three iterations and figure 2.6 illustrates the latest one which is also the most generic version of the three. Figure 2.7 illustrates the second iteration of the model and presents its explicit mapping to the Ellis model of HIB. All three iterations of the model are tightly-coupled and an evolution of the same concept (Fisher and Julien, 2005). They aim to provide the HIB research with a linking framework that will allow the binding of the core HIB and Psychological theories into a concise framework for IR and IS&R research and development.

The findings of Wilson's work (Wilson, 1999) led him to the conclusion that all of the disciplines of widely understood Information Science are indeed well nested and tightly related to each other. This chapter, along with the rest of this thesis aims to illustrate that the generalisation made by Wilson can be taken even further to include the generic research on human behaviour (as illustrated in figure 2.6).

Wilson's contribution to the discipline undoubtedly lies in the integrative view his meta-models provide to the discipline linking various historically older models. Similarly he champions the notion of Information Search Behaviour and IR research being strongly coupled with more generic HIB research and consistency of the mental processes employed. He also acknowledges the complexity of those processes by integrating a significant amount of both external as well as user related factors impacting IR in his latest model (2.6).

Similarly to the models presented earlier the work published by Wilson provides only as set of meta-models and meta-theories and does not effectively address the question of how to cope with the almost infinite complexity of the proposed research paradigm. It does not answer the question of how research or engineering effort is supposed to cope with almost infinite number of combinations and dynamics of the identified factors as illustrated in figure 2.6. If all of the factors as listed by Wilson significantly affect human actions will it be possible to ever predict their outcomes considering their ever changing nature?

This image has been removed

Figure 2.7: Second Wilson's model - Second Wilson's model of HIB and its mapping to Ellis model (Wilson, 1999)

2.3.1.5 Saracevic's model of stratified interaction

Sracevic's stratified model of interaction (1997*a*) unifies the perception of the IR process representing the human aspects of information processing in the same way as the system information processing. Such an approach, even though it can be criticised for the dehumanisation of the user is very valuable for research design and development. It points out the main mechanisms that act during HIB and their interaction also illustrates the flow of information between the components. In the model it was pointed out that the user traits, context and problem at hand are not the only sources of behavioural moderation. As well as this, the properties of the IR system also affect user interaction and their impact is significant. The information, as well as the interfaces, shape our perception and capabilities regarding the information world, and moderate our requirements for IR systems. This is clearly illustrated by the empirical research discussing the information order effects (Buda and Zhang, 2000; Elsenberg and Barr, 1988,?; Huang and Wang, 2004; Joachims et al., 2005; Mantonakis et al., 2009; Murphy et al., 2006) and their impact on perception of information relevance. Even more fundamental effects are introduced by new user interfaces, search approaches or information types. An example of such effects is the introduction of the map based user interfaces as used by Google Maps. They not only change the way people interact with information but also change the perception of the entire search paradigm affecting the user perception of the situation, perception of the cognitive effort required and the problem-solving approach. It also influences the cognitive load and difficulty associated with the problem, which is also related to personality traits (such as the ability to cope with maps) and finally the query representation. Figure 2.8 illustrates the components of the stratified model and the way they interact on various levels of information.

One of the most important aspects of the model is the fact that it underlines the importance of human machine interaction/dialogue and highlights the components that have to communicate in order for the system to successfully work.

2.3.1.6 Bates' Berry picking

The HIB model proposed by Bates Bates (1989) differs from the classical approaches in two ways:

This image has been removed

Figure 2.8: Saracevic's model - Stratified interaction (Saracevic, 1997*a*)

- First of all it notices that the user information need is far from being static and well-defined. The information captured throughout the search process continuously reshapes the searcher's perception of the problem as a consequence affecting both the evaluation of relevance, as well as the search query formulation strategy. This evolution of information need (as illustrated in figure 2.9) is far from being always systematic or well-defined. On the contrary the information need can shift into unanticipated directions. Original goals can be fulfilled, postponed or abandoned in the light of new information.

- Secondly Bates makes it clearer that the information need (or an information Gap or ASOK) is in a majority of cases not satisfied by a single query and a static set of documents. On the contrary the transforming information need is satisfied by individual entities of information found along the way.

This image has been removed

Figure 2.9: Bates' Berry picking - A Berry-picking evolving search (Bates, 1989)

The observations made by Bates are not only well supported by multiple empirical studies, and are consistent with the majority of psychological theories on information acquisition, but what is more important, with the current practice of Internet content design.

From the perspective of this research, and from the perspective of human information retrieval consistency, the work by Bates adds an additional level of complexity. How can one talk about the repeatability of HIB if the information needs are ever changing and if the change is triggered by interaction with individual entities of information? How is a computer system to predict user future information needs, if their evolution is by nature temporal?

Bates (1989) proposes the answer to the above questions. in the guidelines given regarding system and user interface design. Implementation of the guidelines can not only support the user in the berry-picking process but also compensate for system inefficiencies.

2.3.1.7 Principle of Least Effort and Information Foraging Theory

The berry-picking model of HIB proposed by Bates is consistent with other well established information theories. The two theories described below shed a new light on the problem of HIB as they suggest that users will deliberately make sub-optimal or locally optimal decisions about relevance and information interaction. The rationale for the behaviour is suggested to be deeply embedded in the human cognition and be related to the inability of human beings to objectively define the goals of their actions, success metrics and the very high cost of achieving optimal results, which are not compensated for proportionally by the reward.

The key two theories that require attention in the context of HIB consistency and identification of factors moderating this behaviours include:

- Principle of Least Effort (Fisher and Julien, 2005; Zipf, 1949) - It is a well established fact supported by countless empirical studies that people "invest little in seeking information, preferring easy-to-use, accessible sources to sources of known high quality that are less easy to use or less accessible" (Fisher and Julien, 2005)

- Information Foraging (Pirolli and Card, 1999) - Similarly, when searching for information people will try to take advantage of their most primitive instincts developed and optimised for gathering food. They will follow frequently unfounded information scents. They will "berry-pick" following the trail of information and

> actively search for further trails of information. As a consequence they will make decision not based on facts but on incomplete knowledge, employing behavioural heuristics rather than well founded logic.

From this perspective the moderating-factor approach as discussed in this section becomes even more complex, as in order to make the approach complete, the consideration of information uncertainty and the cognitive cost of action would have to be added to the model. This does not only increase the set of moderating-variables but also increases the size of the problem, exponentially affecting all other variables discussed in the previous sections. Moreover, since the phenomena is not widely understood, there exists no research discussing how the phenomena could be embedded into a successful personalised/contextualised IR system design.

The sections above raise a very important question for the problem of human behavioural consistency, namely:

> "Does it matter that in some instances the HIB varies significantly? Is it possible to shift the focus of analysis to task and investigate human consistencies only throughout the tasks that impose behavioural constraints and as a consequence impose certain HIB?"

Further subsections of this chapter focus on the question stated above.

2.3.1.8 Cognitive Work Analysis

The Cognitive Work Analysis approach for IR proposed by Fidel et al. (2004) is a special case of the general Cognitive Work Analysis designed for a wider computer related research and design (Vicente, 1999). Cognitive Work Analysis in a wider context is a "systematic method for establishing the human-system integration requirements for a work domain" (Vicente, 1999). As a consequence the model discussed by Fidel is a systematic method for establishing the human-IR system integration requirements for a work domain.

In contrast to the heavily user-focused approaches to IR, where the user is treated as a black box and a focal point of the system analysis, this approach shifts the emphasis towards a task. The supporters of the approach believe that the work situation and work-specific challenges impose significant constraints on the information system user, reducing the variability of the majority of the variables discussed in the previous chapter. They believe that by designing task-focused system the complexity of human variability as discussed in previous sections can be controlled and in fact the majority of those variables can be ignored.

This model is similar to the Holistic Cognitive Framework and some other models discussed in this section. The model also places the user characteristics, perceptions and activities in the centre of the analysis. Its core assumption is also that the human being is significantly goal driven and that the repeatable goal properties will incur reputable user responses and perceptions of information relevance. The main difference between the models is the generality of the Cognitive Work Analysis approach to other computer science disciplines and the much longer history of use and applicability of the approach.

This image has been removed

Figure 2.10: Cognitive Work Analysis - Cognitive Work Analysis model (Fidel et al., 2004)

And indeed, as illustrated in figure 2.10, the Cognitive Work Analysis focuses primarily on the identification of constraints imposed by the various dimensions of the work/task environment as illustrated by the circles in the diagram. The research focusing on Cognitive Work Analysis correctly notices that the span of tasks within a certain organisation or work environment will not be infinite and will not be uniformly distributed. Similarly the distribution of information worker characteristics will not be uniformly distributed either (due to the selection during the recruitment process and work-related requirements). As a consequence the cognitive work analysis provides a unique framework for reducing the complexity of the moderating-variable factors for the price of the generalisation of the research results (as they cannot be easily transferred to other domains). By shifting the focus from the user and the variability to the key properties of the task that the computer system is supposed to facilitate, the designer can overcome his inability to predict and interpret every users' action and focus on the interpretation of the behaviour within the boundaries of the task and its setting.

Cognitive Work Analysis provides a multi-stage, very well-defined and mature framework for solving well-defined system design problems what are also transferable to the IS&R domain. Cognitive Work Analysis as an approach is a tool for system design and not a platform for understanding the impact of individual differences on system performance. Regardless it has proved to be very effective as a system engineering utility and has currently hundreds of real life applications and still is an active area of research within the Human Factors and related domains.

The work proposed by Fidel et al. (2004) is as a consequence perfectly valid for a set of IS&R problems where the task is recurring, and can be clearly captured and defined. For those problems it provides a good framework for coping with complex system complexity (Jenkins, 2009).

From the perspective of this research and the problem of moderating-factors and their impact on behaviour the proposed approach does not suggest that the number of moderating-factors is small or that they do not play a role in the shaping of HIB. Instead the approach is based on the assumption that in certain controlled environments those variables change very little and as a consequence have a limited impact on the variance of the observed behaviour.

The primary difficulty in the application of the Cognitive Work Analysis approach to the HIB is its dynamic nature and the wide span of problems to which it is being applied. The greatest value of IS&R utilities lies in the end in the support of variable and ill-defined problems. Secondly, even though the Cognitive Work Analysis is very established in the HIB domain there exists a very limited amount of literature within the HIB and IS&R making its application to the problem of relevance-feedback or collaborative-filtering non trivial.

2.3.1.9 Task-focused approaches

A significant proportion of the modern IS&R approaches ignore the notion of task during the research design and implementation. This seems to be a futile effort as there exists a compelling body of evidence illustrating that it is not only the task and its complexity but also other factors, including the stage of the task which significantly affect the perception of relevance, as well as the criteria for the selection of information channels and objects Vakkari (2008). The empirical results strongly suggest that task and its properties are a major factor affecting the HIB and as a consequence the design of IS&R utilities.

Advocates of the task-focused approach to HIB believe that our knowledge of the search task is incomplete without the knowledge of the task that triggered it. They believe that in order to gain a complete picture of the HIB we need to further understand the interactions between the search and work task as well as the fundamental properties of the work task.

The two most widely recognised work-task focused models/approaches in IS&R research were introduced by Kuhlthau (1991) and Vakkari (2003) as well as Bystrom and Jarvelin (1995); Campbell (1988); Sutcliffe and Ennis (1998) and Sutcliffe et al. (2000). Their work can also be perceived as a moderating-factors approach to HIB. The model proposed initially by Kuhlthau (1991) and then further reviewed by Vakkari (2003) introduces a methodology for breaking down the work tasks into conceptually separate stages which then can become the unit of analysis. They observe that the stage of the task affects the search process and perception of relevance and as a consequence must be a part of the research design. The conceptual comparison of the approaches is illustrated in figure 2.11.

It is important to note that the above discussed work is far from being an ethnomethodological approach to IS&R system design. On the contrary, it is very deeply founded in the cognitive/black box model paradigm of research. From the perspective of this research and the concept of factors moderating HIB, the properties and stages of task as listed above, become yet another group of dimensions in the complex

This table has been removed

Figure 2.11: Task/Activity focused approaches - Comparison of the key concepts of the work carried out by Kuhlthau and Vakkari (Vakkari, 2003)

space/black box of human cognition. The task-focused approached to IS&R do not make the task the centre of analysis but treat it as an additional set of input variables that needs to be taken into account during research design.

2.3.1.10 Holistic Cognitive Framework for IR

Ingwersens' Holistic Cognitive Framework for IS&R claims to be an integration of the main IS&R models and research frameworks (Ingwersen and Järvelin, 2007) as discussed in the sub-sections above. It promises to be a holistic platform integrating both user as well as task-focused approaches discussed earlier.

The components of the model are listed in figure 2.12. The framework takes the cognitive approach to the problem of IS&R research and places the Cognitive Actor (i.e. the user) in the centre of the research and design process but not ignoring the task and its interaction context either. It acknowledges that all other components of the model are inevitably only cognitive and subjective perceptions which are being created by the user but also appreciates the impact the task/action has on that perception. Intentionally the key attributes of the "Cognitive Actor" relate to the work and search task. Such a central position of the work and search task is not only grounded in the cognitive approach itself, but also on the observation that human beings are significantly goal driven, which is discussed in the next chapter along with the "Perceptual Control Theory".

The model also places an emphasis on the importance of Context in the information interaction process and differentiates three main categories of context: Organisational, Social and Cultural. Neither of the other components of the model can stand alone and they are all deeply placed in context and cannot exist outside of it. The actor cannot operate without context and is never stripped of his organisational, social and cultural background. Similarly the search systems and user interfaces cannot function without a context as they are being created for a particular purpose and a clearly defined group of people. The components of the model are connected by either unidirectional or bidirectional arrows indicating the direction of influence.

The model aims to challenge the traditional Laboratory approach to IS&R research and proposes a new and concise research framework for user-centred design and development. It provides a means for controlling some of the moderating-factors affecting

This image has been removed

Figure 2.12: Cognitive Framework - Holistic Cognitive Framework for IR (Ingwersen and Järvelin, 2005*b*)

the research result validity and repeatability under different settings at the same time controlling the problem complexity (Ingwersen and Järvelin, 2007).

As one can see the Holistic Cognitive Framework for IS&R is also a moderating-variable approach, listing 9 main dimension factor groups which have to be resolved to appropriate moderating-factors on a per problem basis, depending on the tasks carried out by the user, context and user groups. The approach to the selection of appropriate moderating-factors and the guidelines regarding their amount and quantification are also available with the model (Ingwersen and Järvelin, 2005*b*).

It is important to note though that as the name suggests the proposed framework is very much a cognitive approach to the problem and as a consequence suffers from all of its limitations. The problem of cognitive approach limitations and the overall complexity of the moderating-factors approach is being discussed in the next chapter.

2.3.2 Moderating-factors: Cognitive Psychology

The moderating-factors approach to the problem of prediction of behaviour is very appealing both in the domain of psychology, HIB and IS&R as it follows the behaviouristic approach. By introducing the moderating-variables the human being can be represented as a "Black Box", without jeopardising the notion of attitude-behaviour consistency which is essential for a system-centred research like that. Owing to the introduction of moderating-factors, or in IR literature the "context", the question "whether attitudes guide behaviour", or explicit feedback is related to future information behaviour as stated by Wicker (1969), could be abandoned. Owing to the rich set of examples illustrating high correlation levels between attitudes and observed behaviour (Goodmonson and Glaudin, 1971; Kelley and Mirer, 1974; Seligman and Miller, 1979), a new question was formulated, namely, "under what conditions attitudes guide behaviour?". The question re-appeared frequently in literature and was discussed very extensively by Fazio asking, "Under what conditions do what kinds of attitudes held by what kinds of individuals predict what kinds of behaviour" (Fazio and Zanna, 1981). The same question reoccurs in the HIB and IS&R research in a changed format adjusted to the information interaction problem:

- HIB: Under what conditions, do what kind of information-related attitudes, held by what kind of information users, predict what kind of information behaviour?

- IS&R: Under what conditions can what kind of information behaviour, held by what kind of information users, predict what kind of information attitudes, for the further use for information recommendations?

The moderating-factor approach was very successful in finding the answer to the above mentioned question in cognitive psychology. The moderating-factor models such as MODE (Fazio, 1990), have managed to provide the research community with a framework to work systematically with the moderating-variables and led to a deeper understanding of how the processes which control the dispositions of humans affect their actions. Throughout the decades a 'rather lengthy catalogue' (Fazio, 1990) of moderating-variables was identified significantly extending the understanding of when moderating-factors guide behaviour and when they do not (Fazio and Towles-Schwen, 1999). The literature did not manage to answer the third generation question (Albarracn et al., 2005) of 'how' the relationship between the attitudes and behaviours is controlled. A great review by Cooper and Croyle (1984) points out that there exists no consistent theory that allows an answer to this question. As a consequence, it was not possible then to make higher level generalisations regarding the impact of moderating-factors on behaviour and no consistent research methodologies existed. The situation is changing currently and our understanding of the attitude-behaviour consistency mechanisms is improving (Albarracn et al., 2005).

2.4 Conclusions

The literature review carried out in this chapter clearly illustrates that it is very difficult to claim consistency of human behaviour and HIB. Some even claim that individual instances of human behaviour are inconsistent (Ajzen, 2005) and that it is not safe to assume otherwise. This finding is a significant challenge for the contextualisation and personalisation of IR. Even though moderating-factor theories promise a mechanism for potential resolutions of the problem they are very limited in their ability to address consistency. The review illustrates that the variability of human behaviour related to moderating-factors is not marginal but affects the key decision-making processes. This affects our perception of information relevance. It also illustrates the complexity of the problem by identifying an overwhelming amount of factors that affect human behaviour and exponentially more interactions between the factors. Moreover the comparison of HIB and IS&R results with the cognitive psychology, showed a high similarity between the mechanisms responsible for decision-making. This also revealed that even more factors are to be considered in HIB prediction which will further increase the complexity of the problem.

Even though there exists a very mature research on behavioural moderation and a vast amount of models aiming to describe HIB their ability to support commercial development of personalised and contextualised systems is very limited. This creates a need for new research methodologies and research frameworks.

Chapter 2: Summary

The problem of human behaviour consistency as well as human information behaviour consistency is not clearly defined nor understood in the modern literature. The extent to which human beings are consistent/repeatable/rational is unknown and varies significantly with a multitude of factors.
The problem of human information behaviour inconsistency is observed from the first days of IS&R research. The early laboratory studies using more than one expert group to assess relevance identified significant discrepancies in relevance estimation and only a small proportion of documents was assessed as relevant by distinct groups.
This issue was diminished for decades because of the conclusions of the study by Lesk and Salton (1968). They hypothesise that the variation of user relevance does not impact significantly the efficiency of the IR algorithm employed or the evaluation process. This conclusion is not applicable for contextualised/personalised IR solutions as they are intended for particular individuals in context and not for an entire IR population.
This lacks of human consistency is also observed in psychological literature. More than 80 years of cognitive research illustrated significant variability of human decision-making process and a lack of statistically significant correlation between individual instances of behaviour.
This issue is frequently addressed through the introduction of moderating-variables into the experimental design which intend to increase the predictive validity of the experiments. Factors such as age, sex, personality, domain expertise and many more were used to explain this observed human inconsistency.
A significant majority of modern HIB models and IS&R approaches are based on the idea of the moderating-factor. Models such as Sense-Making Theory, Wilson's model, ELIS and many other reviewed in this chapter list an enormous number of moderating-factors that affect user information choices.

Table 2.2: Summary of chapter 2 part 1

Chapter 2: Summary

The moderating-factor approach has a number of limitations, the main one being the share quantity of the moderating-factors that have a statistically significant impact on behaviour which makes the approach unmanageable. Moreover, the majority of the factors are proven to interact with each other and have a very selective nature further increasing the complexity of the approach. The moderating-factor approach does not explain the cognitive mechanism responsible for choice and as a consequence is very difficult to use for the high accuracy prediction of information relevance.
Overall the review illustrates that human decisions have a significant potential for being consistent/rational. It showed that the complexity and also lack of understanding of the human decision-making process makes the use of this consistency very difficult. It revealed the fact that the moderating-factors approach is not appropriate for this type of a problem and that other approaches for behaviour predictions have to be employed in order to deliver business ready IR solutions.

Table 2.3: Summary of chapter 2 part 2

3

General Black Box Model

3.1 Introduction

The previous chapter has discussed a wide range of factors which impact the HIB. This also includes the perception of information relevance which was shown to change significantly between the users. The review of HIB research was carried out to pinpoint the factors which affect the human decision-making process. Importantly it illustrated the impact of moderating-factors on user decisions. Additionally the analysis of the cognitive literature showed the similarity of the cognitive processes used in everyday and information-related decision-making process.

This chapter builds on the understanding of the role of moderating-factors and discusses a simplified model of HIB referred to in this thesis as the Black Box model of HIB. The Black Box model is only concerned with the input and output variables of humans and intends to ignore the existence of human cognitive processes. This allows IR research to be described in a very consistent way and identifies the limitations of the moderating-factor approaches discussed in the previous chapter. The Black Box model is then further used in the next chapter to discuss the problem of human consistency and validate the proposed research methodology.

3.2 Backbone of the model: moderating-factors

As discussed in the previous chapter the prediction of HIB based on moderating-factors is non-trivial. A countless number of moderating-factors involved in the human decision-making process and the complex relations between them make modelling and prediction of future HIB very prone to errors. The interaction between moderating-variables is frequently neglected in the IR literature and its impact on HIB is rarely analysed. This is a significant issue as the impact of variable interactions on HIB is non-trivial. There exist countless examples revealing the complexity of the problem. One of the most characteristic examples illustrating the issue is the impact of moderating-factors on user efficiency/efficacy.

User efficiency during IR is defined by the HIB research in various ways. It is frequently defined as:

- The total amount of time required to complete the entire IR task.
- The average amount of time spend in a search session or on a part of the main task.
- The number of relevant elements of information found in a period of time.
- The number of relevant elements of information found in comparison to all of the information available.

The differences in the way the efficiency is defined and measured make the investigation of factor interactions impossible. Therefore for the purpose of this illustration a single definition was selected (see Kim (2001*b*) for details). The key factors which are widely understood to have a significant effect on user HIB efficiency are as follows:

1. Domain expertise

   The first parameter that has a statistically significant effect on user IR efficiency is the user's domain expertise. It is widely understood that the relation between the domain expertise and human efficiency is governed by the power law of learning (Anderson, 1982; Anderson et al., 2004, 1985; Newell and Rosenbloom, 1981; Rabbitt and Banerji, 1989; Seibel, 1963) as shown in Figure 3.1a. The relationship between domain expertise and efficiency is significantly moderated by the user cognitive style (Witkin and Berry, 1975; Witkin and Goodenough, 1981; Witkin et al., 1980) and web expertise (Hirsh, 2004; Hölscher and Strube, 2000; Kim, 2005; White and Drucker, 2007; White et al., 2009; White and Horvitz, 2009; Wildemuth, 2004; Zhang et al., 2005) which is illustrated in Figure 3.1b.

These images have been removed

Figure 3.1: Domain Expertise VS Efficiency - a) Efficiency vs Familiarity level (Kelly and Cool, 2002); b) Efficiency vs search expertise and Cognitive style (- Field Dependent Participants, - Field Independent Participants) (Kim, 2001*b*)

2. Task type

   The impact of task on user efficiency and also HIB efficiency is widely recognised in both psychological (Engel et al., 1999; Greeno, 1978; Voss and Post, 1988) and IR literature (Gwizdka, 2008; Kim, 2006, 2008*b*; Kim and Allen, 2002; Li and Belkin, 2008; Tu et al., 2008; Turpin and Scholer, 2006). The IS&R research recognises that task and its parameters such as complexity (Gwizdka, 2008); facet (Li and Belkin, 2008); definition and structure (MacGregor et al., 2001; Simon, 1973; Voss and Post, 1988); and psychological classification (Engel et al., 1999; Greeno, 1978); have a significant impact on the IR process. Moreover the research recognises that the task interacts with a variety of factors such as epistemological beliefs, cognitive style, domain and search expertise and many others (Tu et al., 2008).

3. Epistemological beliefs versus user efficiency

   Additionally the efficiency of the user is significantly impacted by epistemological beliefs (Tu et al., 2008) as shown in figure 3.2. The results suggest that participants with more constructivist-oriented epistemological beliefs tended to visit fewer pages but explore more deeply than the rest of the students. Overall they are better in open-ended questions as they have the ability to pay more attention when filtering information and overall, achieve better search outcomes.

This imnage has been removed

Figure 3.2: Epistemological beliefs VS Behaviour - Correlations among behavioural variables epistemological beliefs, student background and tasks (Tu et al., 2008)

> Figure 3.3 presents the regression models that can be used to predict the performance of students performing open and closed-ended tasks (White and Roth, 2009). The results reveal that for the closed-ended tasks the student web-experience was the primary factor deciding user performance. On the other hand, for the open-ended tasks, the epistemological beliefs and meta-cognitive abilities are the main primary, and only factors, affecting user success.

Based on the three factors discussed above which impact user HIB efficiency, the plot presented in Figure 3.4 was created. The figure shows how 3 moderating-factors (domain/web expertise, task type and epistemological believes) affect user efficiency. As shown in the figure the factors are not linearly independent. The figure shows the existence of significant interactions between the factors. In order to understand the impact of one factor on user efficiency, the other two factors have to be captured and included in the analysis, to avoid a significant error. Moreover the illustrated interactions are highly non-linear, which further increases the complexity of the problem.

These three factors contribute to a small proportion of user-efficiency variance. As illustrated in the previous chapter HIB is highly affected by a multitude of other factors, such as age; presentation order; other task and cognitive style parameters; language; culture and context of interaction (Kim, 2001*a*). Without further evidence

Figure 3.3: Task VS Behaviour - Regression models for predicting user performance across different tasks (White and Roth, 2009))

These images have been removed

Figure 3.4: Efficiency prediction - Interpolation of the Efficiency as a function of Web Expertise, Epistemological Beliefs and Task Type (Open Ended =1, Close Ended = 0) (Tu et al., 2008). Hue represents the efficiency score.)

it is difficult to assume that those additional factors will be independent, making the problem simpler. On the contrary, empirical evidence exists which illustrate further interactions between the parameters such as age/gender (Almutairi, 2011; Large et al., 2002; Roy and Chi, 2003; Weiser, 2000), user domain/web expertise (Marchionini et al., 1990; White and Drucker, 2007) and multiple other factors (Ford et al., 2005; Kralisch and Berendt, 2004; Li and Kirkup, 2007; Nisbet and Norenzaya, 2004).

The modern HIB models do not provide the tools to model complex interactions between moderating-factors. Therefore the selection of a model like a Black Box model is needed in order to simplify the process. The next section proposes a Generic Black Box model.

3.3 General (Black Box) Model of HIB

The proposed General (Black Box) model is inspired by two models:

- The behavioural "Black Box" model (Holland and Skinner, 1961; Skinner, 1938).
- The cognitive "Perceptual Control Theory" (Forssell and Powers, 2009).

The proposed model provides a vocabulary and a thinking-tool to deal with complex HIB problems. The model allows for systematic discussion, research design and result analysis of any research concerning moderating-factors. Importantly it provides a wide range of mathematical and statistical tools that can be used for result analysis and the comparison of results with existing research.

The proposed model was also heavily inspired by the Rijsbergen's (2004) Geometry of IR. Rijsbergen proposes a generic geometric framework that allows for formalisation and generalisation of multiple IR approaches such as logical, probabilistic and vector space models (Rijsbergen, 2004). The framework helps to formalise several behaviour based IR approaches such as "Pseudo-Relevance Feedback", "Relevance-Feedback" and "Ostensive Retrieval". The proposed Black Box model serves a similar purpose in HIB research. It defines a simple and consistent framework that allows for the description of HIB research results. It allows the HIB and behavioural/cognitive IR to use the same vector space vocabulary and discuss the research related concepts and issues in a consistent and highly formalised manner.

The subsequent section discusses the existing models, which provide an insight into the proposed general black box model.

3.3.1 Existing models

3.3.1.1 Behavioural Black Box

The Black Box metaphor is frequently used by radical behaviourists (Holland and Skinner, 1961). The idea of radical behaviourism is based on the assumption that human behaviour can be studied as a natural science. Radical behaviourists perceive human responses as pre-determined and humans not having the element of free will. Such an approach to the problem of behaviour analysis allowed behaviourists to treat the human cognitive and learning processes as a "Black Box" device as shown in Figure 3.5.

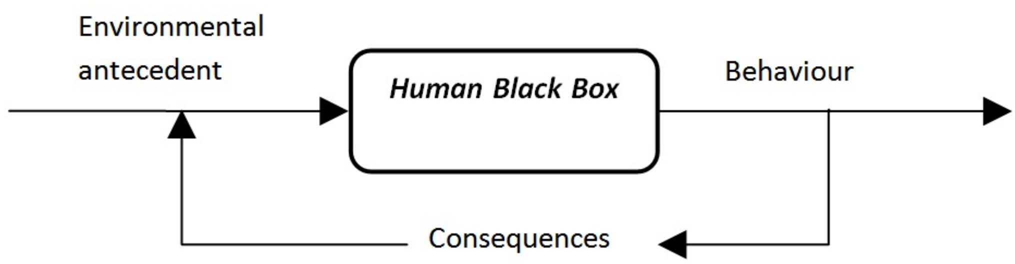


Figure 3.5: Behavioural Black Box - Behavioural Black Box Model

Such simplified approach, even though often criticised, is still frequently used in psychology and multiple other domains. Sandhusen (2008), for example, used the approach to predict customer purchasing behaviour.

One of the most widely recognised critics of the radical behaviourism approach and also the behavioural black box model is Chomsky (1959) who proposed his own Black Box model metaphor (also referred to as a generic "Mind"). When introduced in the late 1950s it allowed for the first time, to approach the problem of language and language development from the holistic perspective. The proposed Language Acquisition Device (LAD) and Transformational-Generative Grammar (TGG) allowed for the integration of biological, mental, social, and cultural factors affecting human language and cognitive development.

3.3.1.2 Perceptual Control Theory

The "Perceptual Control Theory" (Forssell and Powers, 2009), as shown in Figure 3.6, demonstrates that humans are not automata responding unconditionally to external sensory input. Rather, they are purposeful beings planning their behaviour and controlling their perception of the external environment. The visual model of the theory (Figure 3.6) illustrates that humans, when faced with a problem, purpose-of-action or goal, behave like an engineering control system (Figure 3.7), providing the surrounding environment with a negative feedback. The model is very useful in the context of HIB analysis as most of the current research is heavily user-task and search-task focused and as a consequence it can be very easily interpreted using the model.

The goal of negative feedback, provided through mechanical or verbal response is to change the environment to achieve or maintain the desired goal. An important observation of the theory is that the user is not controlling the environmental variables directly but instead is processing internal perceptions of both the surrounding world and the actions which follow as a result of those internal perceptions. The Input and Output elements of Figure 3.6 are responsible for the interpretation of the external world and the translation of the perceptions of actions into feedback affecting the environment. As a consequence, the 'Comparator' element is not driven by the objective sensory input and objectively defined task or user goal. Instead, it is provided with the interpreted, limited and subjective information about the outside world and the desired outcomes of actions. Moreover the responses created by the "Comparator" element are not directly related to user-physical responses but are just perceptions of actions to be undertaken. Only the "Output" element translates them into actions. As a result the response of humans will be significantly affected by them, and their ability to interpret the external world, make effective decisions and translate them into actions.

This ability of the user relates to multiple user traits such as user personality, knowledge, experience, character, cognitive style and many others. In the further discussion this trait affecting the user cognitive process will be referred to as the User State.

"Perceptual Control Theory" is a generic cognitive theory which is also valid to address the problem of IR and HIB. The only difference being the scope and dynamics of the reference goal used to control the interaction with the outside world. When

This image has been removed

Figure 3.6: Perceptual Control Theory - The essence of the "Perceptual Control Theory" (Forssell and Powers, 2009)

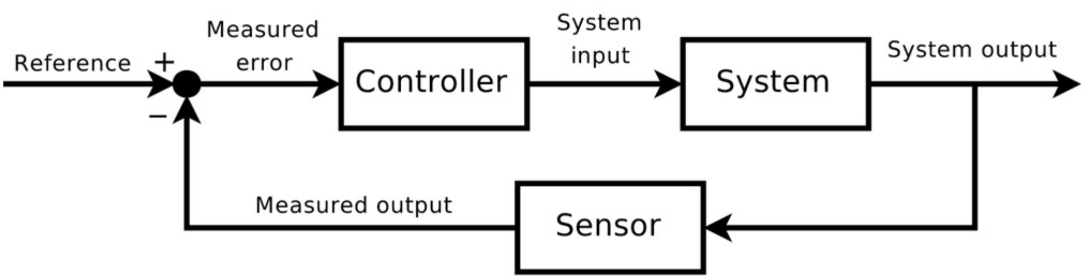


Figure 3.7: Classical control system - Classical control system

the user attempts to solve an information-related task he focuses his sensory input on the available information-related resources and interprets them creating the subjective representation/perception of the information that is being perceived. The process of information interpretation is highly personal and highly dependent on the above defined User State. Figure 3.8 reflects this observation by defining the "Input Interpreter" as a function of not only "Sensory Input" but also "User State". The next cognitive step involves the "Comparator" which compares the current "Perception of the Environment" (i.e. in the case of HIB, with the currently available information) with the requirements of the information-related task at hand. Similarly, the comparison process that leads to the subjective perception of difference is not exclusively a function of a task and environmental perception. It will be highly dependent on the User State leading to different perceptions of information, depending on User State variables such as user experience, knowledge, personal character, etc. Finally, the response to the Perception of Difference is also moderated by the User State, significantly varying user information behaviour. Figure 3.8 presents the above mentioned information-related variables and the relations between them, demonstrating the importance of the User State.

The figure above (Figure 3.8) presents a variation of the "Perceptual Control Theory" adjusted for the purpose of HIB analysis. It also underlines the role of User State and the functional nature of the components taking part in the cognitive process. This functional nature of the components of the above mentioned model is the core of the Black Box model discussed in the next section.

3.3.2 Proposed General Black Box Model of HIB

"Perceptual Control Theory" (Figure 3.8) assumes that the user control-system can be represented as a probability-density function, otherwise there exists no causal relation between the sensory input and the observed behaviour. Importantly this assumption of human non-randomness is a key for development of IR systems and a foundation for further discussion. In other words, it is only possible to develop IR systems if there exists such a function that maps each Perception of the Environment, User State and the Perception of Task, to an instance of Human Information Behaviour (Figure 3.9), or its probability distribution function at that point.

The black box model can be summarised in one of the following expressions:

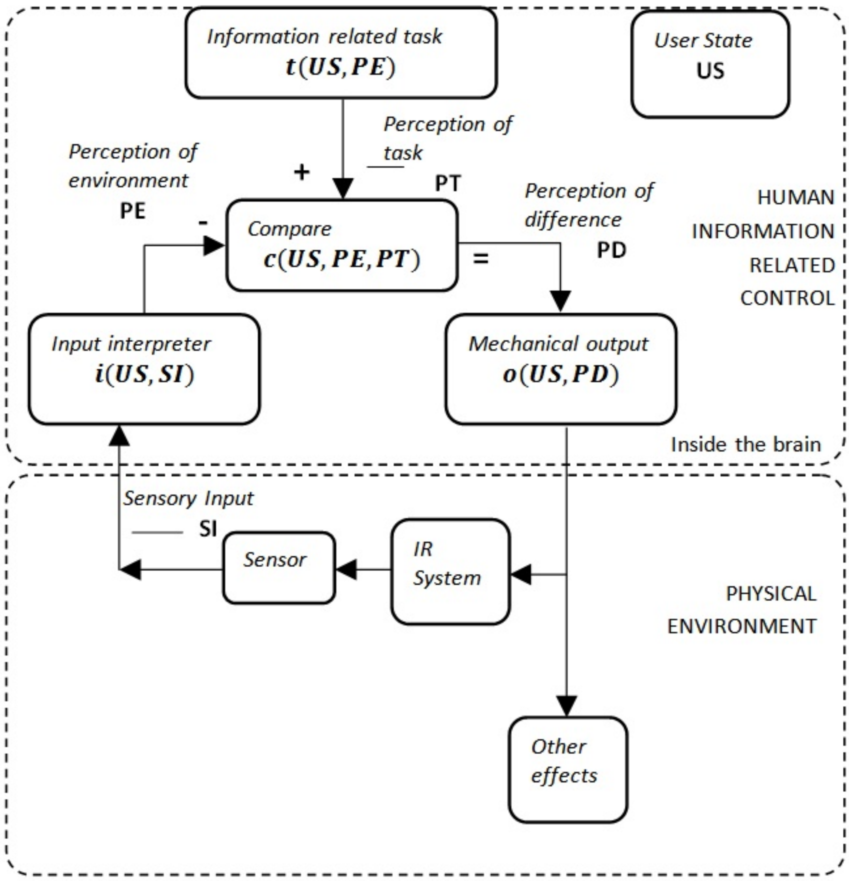


Figure 3.8: Perceptual Control Theory of HIB - HIB specific interpretation of the Perceptual Control Theory

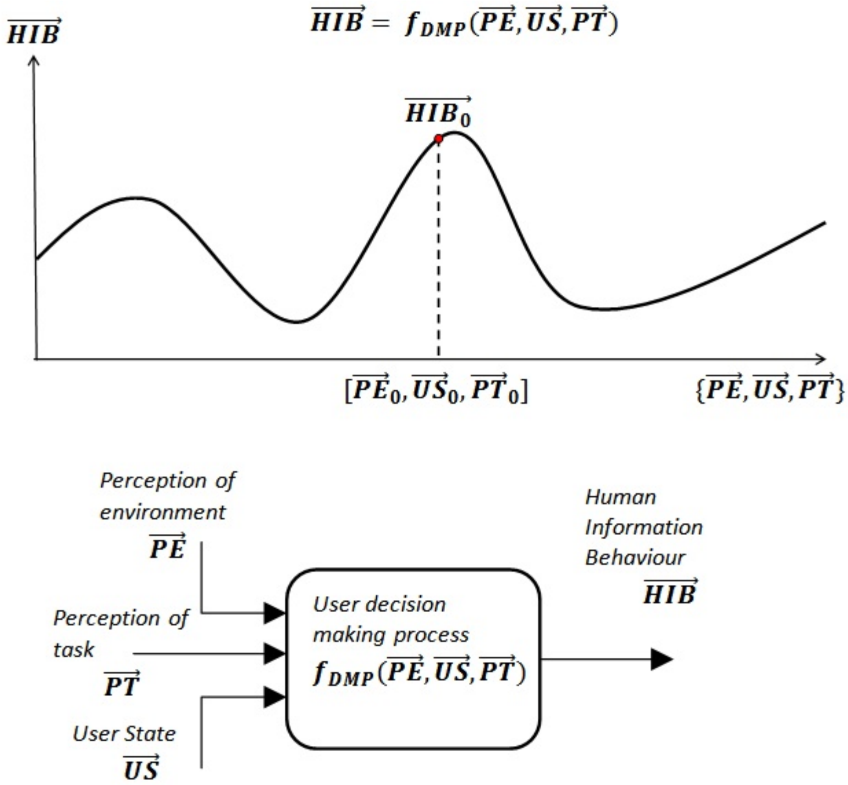


Figure 3.9: General Black Box Model - Black box model of HIB

$$\overrightarrow{HIB} = f_{DMP}(\overrightarrow{PE}, \overrightarrow{US}, \overrightarrow{PT}) \tag{3.1}$$

Figure 3.10: Deterministic Black Box model equations

$$P(\overrightarrow{HIB}) = f_{DMP}(\overrightarrow{PE}, \overrightarrow{US}, \overrightarrow{PT}) \tag{3.2}$$

Figure 3.11: Indeterministic Black Box model equations

- For all possible "Perceptions of the Environment", "User States" and "Perceptions of the Task" to be carried out, there exists a function f_{DMP} of user decision-making process that associates the combination of those three vector variables to a single instance of Human Information Behaviour $\overrightarrow{HIB}$:
- For all possible "Perceptions of the Environment", "User States" and "Perceptions of the Task" to be carried out there exists a function f_{DMP} of user decision-making process that associates the combination of those three vector variables to a single HIB probability density function P($(\overrightarrow{HIB})$

Regardless whether human beings are fully functional and are incapable of making free decisions (point 1) or whether they can make free decisions but their choices are rational (point 2) the black box model holds and can be used for the IS&R purposes. It is worth discussing briefly how human rationality/consistency is manifested mathematically in terms of the probability distribution function. Should the human be irrational, the expected behaviour distribution should be uniform and all instances of behaviour uniquely probable. Rationality introduces deviations from the uniform distribution functions and only certain behaviours are probable. Figure 3.12 illustrates an abstract probability distribution function of a rational behaviour. Every peak corresponds to the most probable behaviour corresponding to a single combination of PE, US and PT.

One can argue, that the assumption of human rationality is incorrect as there are countless examples of human irrational choice. It is important to notice though that the Black Box model does not discuss the objective rationality of human choices, it focuses on human causality from a human-perceptions perspective. The perception of the environment, by definition, can be significantly different from the real/objective shape of the surrounding world (if such a notion exists at all). From this perspective, every action, even if perceived by the surrounding observers as illogical and irrational,

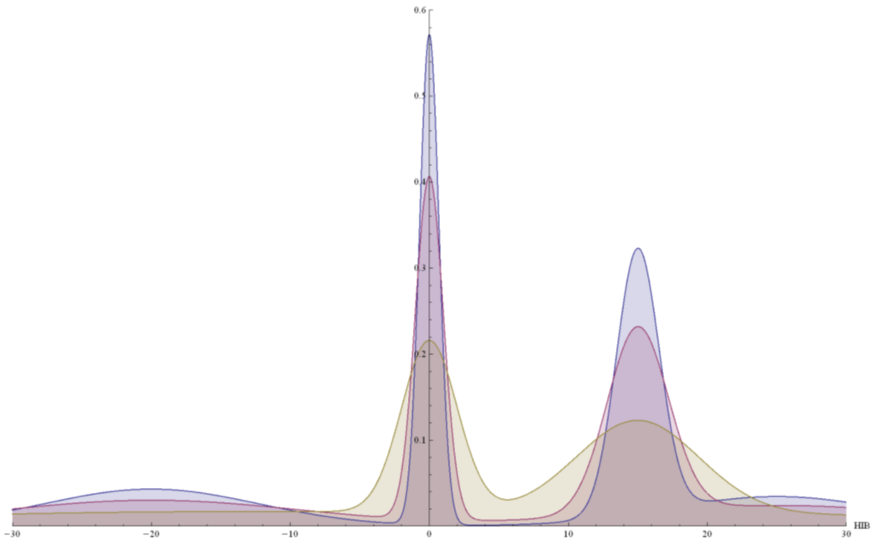


Figure 3.12: Probability of HIB - Visualisation of probability density function of Human Information Behaviour at $P_0(\overrightarrow{HIB}) = f_{DMP}(\overrightarrow{PE_0}, \overrightarrow{US_0}, \overrightarrow{PT_0})$

is a natural and unavoidable consequence of the user internal state and the perception of the surrounding environment.

The model defined in such a way is unique, as it treats the user decision-making process as a black box, not going into the detail of cognitive processes responsible for the shape of the decision-making function f but instead it focuses on:

- Input parameters moderating the HIB which can be classified, following the "Perceptual Control Theory", into three generic vectors: PE - Perception of the Environment, US User State and PT Perception of Task.
- The shape of the Black Box function f, its dynamics and its potential probability distribution.
- The output of the model being the HIB vector and its components.

The Black Box model introduces a level of abstraction that allows for the unification of the moderating-fctors research as discussed in the previous chapter of this work. Every point on the abstract f_{DMP} illustrated in Figure 3.9 corresponds to single instance of HIB that was triggered by a certain combination of PE, US and PT.

In the example in Figure 3.9 the point $[\overrightarrow{PE}_0, \overrightarrow{US}_0, \overrightarrow{PT}_0]$ corresponds to a single instance of Human Information Behaviour $\overrightarrow{HIB}_0$. Alternatively in a case when there exists an element of free will/ behavioural choice/ behavioural randomity the point $[\overrightarrow{PE}_0, \overrightarrow{US}_0, \overrightarrow{PT}_0]$ corresponds to the HIB probability density function (Figure 3.12).

Such representation of the HIB is flexible and abstract enough to allow for the description of all the cognitive processes. This also includes the cognitive processes involved in everyday information behaviour. In contrast to the existing HIB models, the proposed model provides an effective means for the analysis of the observed results, the interpretation of the interaction of the different moderating-factors and their variability in time. Moreover, it allows us to explain the deviations of the observed results from the expected values predicted by the existing models, by the existence of other latent or simply unknown moderating-variables that affect the variation of observed data. Finally, the Black Box model equips us with a variety of tools that can facilitate the analysis of the behavioural results. For example, in the case discussed above, Principle Component Analysis can be employed in an attempt to remove the correlation between the variables (for example those occurring among different cognitive style classifications or among different task classifications) or alternatively, any Factor Analysis technique can be applied to investigate other sources of data variation.

It is important to notice that the Black Box model discussed above takes advantage of user Perceptions of the surrounding world which are significantly moderated by the User State. This issue is discussed in the next section of this chapter.

3.3.3 Application of a Black Box model moderating-factors approach

The Black Box model, similar to the "Perceptual Control Theory", highlights the fact, that the behavioural decisions, also related to information choices, are not made based on sensory input, but are made based on subjective perceptions which are highly dependent on user individual and unique state. Because of the latent nature of the variables of the model the model, is difficult to apply for the development of real life information systems.

The sole goal of an IR System (Figure 3.8) is to process the negative feedback provided by the user, to minimise the Perceived Difference between the required information and the perceived information, by not removing from the user the Perception of Control. This goal is very difficult to achieve because the values of the latent variables (i.e. Perception of the Environment, User State and Perception of Task) cannot be objectively known. Instead, the real life IR system has to take advantage of the measurable variables. The IR system aiming to model HIB, will take advantage of the following variables:

- $\overrightarrow{P}$ Formal representation of a problem to be solved. It is an approximation of the subjective Perception of the Task.
- $\overrightarrow{C}$ Formal representation of a context in which the problem to be solved is being set. It is an approximation of the subjective Perception of the Environment.
- $\overrightarrow{U}$ Limited information about the User State variable.

To differentiate between the variables of the Back Box model and the variables used by the actual model under investigation, is very important during the design and interpretation of HIB and the behavioural IR research. For example, the review from the previous chapter, clearly illustrated how personality and thinking style affects the objectively measured domain expertise and moderates its impact on the user reading-time. Similarly, it is illustrated how the information-presentation order has affected the perception of relevance, and how this impact was further moderated by the user Internet and domain experience. To summarise, the research taking advantage of objective measures of the environment, content and the problem at hand, inevitably introduces an error into the HIB prediction process. The error can be minimised if subjective or implicit measures are used in the research design and implementation. Similarly this also applies to the process of modifying the context of the information interaction and the exact scope of the information problem. For example the research focusing on explicit forms of relevance-feedback, particularly in laboratory conditions often affects the context of interactions. It also imposes tasks on the users affecting user engagement and relevance estimation tolerance. Even the ethnographic approaches to explicit relevance-feedback bring about a change to the context, by introducing additional rating-tools and imposing their use on the users. This has a widely documented impact on the user rating-process which is dominated by the negative feedback (Manning et al., 2008).

Finally for a moderating-variable IR implementation to provide sufficient predictive validity to be used in real life, the moderating-factors analysed by the system would

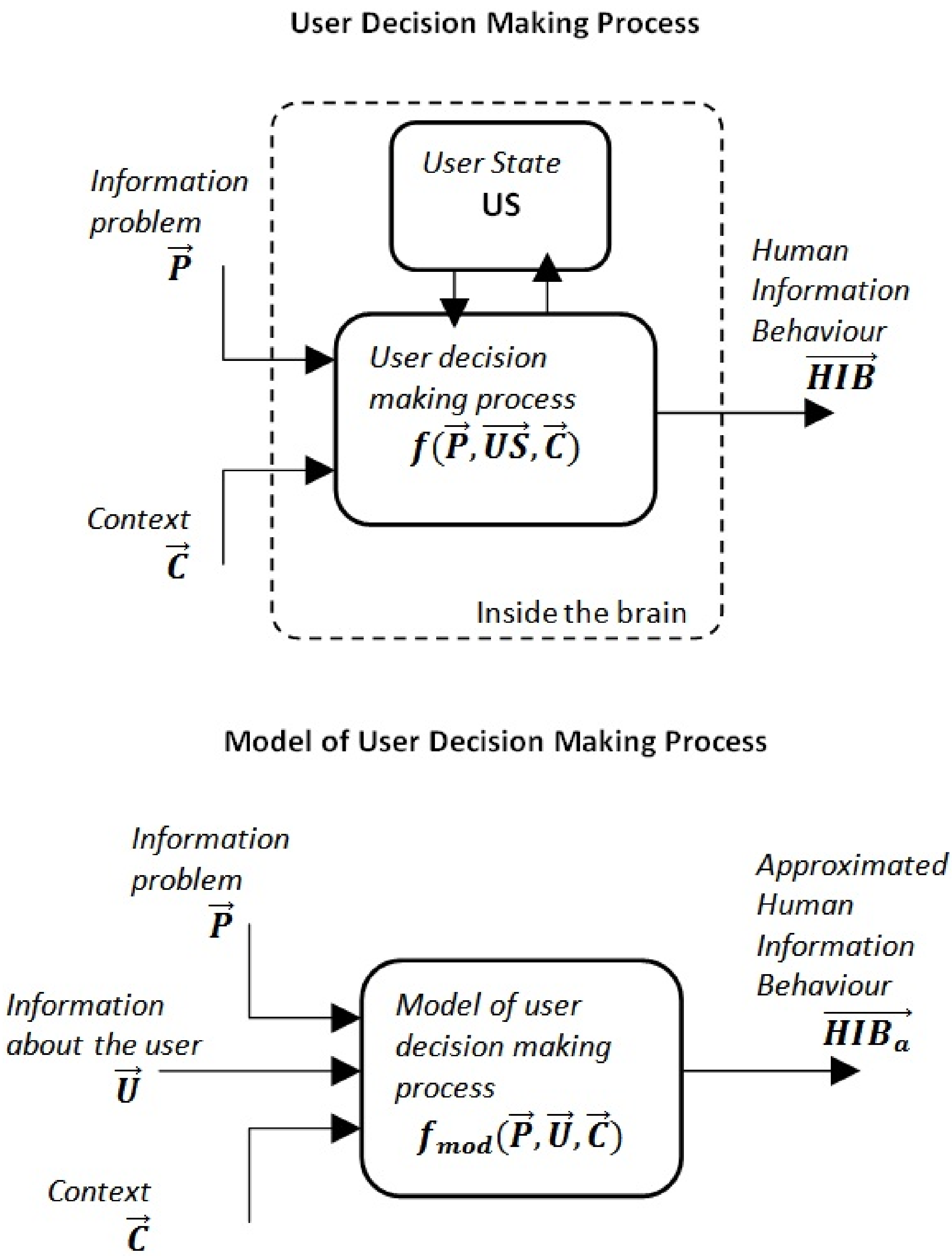


Figure 3.13: Human Cognition VS System Centric Model of Human Cognition - Difference between the user decision-making process and its automated model

have to explain a sufficient proportion of variance to offer the user a non-random search experience. Should the proportion of variance explained be low, users will be receiving significantly varying results, even though from their perspective the context of retrieval might have not changed or changed to any marginal extent. Moreover, the group of selected moderators would have to be independent from other moderators and the non-observed variables should not impact the behaviour at all, or only to a marginal extent.

3.4 Conclusions

This chapter introduced the Black Box Model of HIB along with the theories lying at its foundations. Furthermore the model was used to illustrate the real life challenges of the moderating-factor approaches to IS&R research:

- Our inability to capture subjective perceptions of the environment, user state, and task which are used by the user to make "rational decisions".
- The complexity of explaining high levels of behavioural variance with the limited number of moderating-factors.
- The very selective nature of moderating-factor based prediction making the approach difficult to use, even during the ethnographic study.
- The complexity of high level moderating-factor interaction affecting directly the feasibility of the approach.

The above mentioned challenges mean that even when faced with rational human behaviour our ability to predict it accurately and consistently is low. Those challenges are a significant thread to the moderating-factor based HIB research and the goal of the next chapter is to discuss those challenges further and propose a new innovative approach in order to overcome them. The next chapter proposes a new model that can be used for the HIB research which removes the challenges related to the use of moderating-factors. The proposed model is based on two well established theories (i.e. Holistic Cognitive Theory of IR and Theory of Planned Behaviour) which focus on the notion of rationality/consistency in order to simplify the variability of moderating-factors. The proposed approach takes advantage of their strengths to propose a consistent and simple methodology for IS&R research (Integrated Framework for HIB in-situ).

Chapter 3: Summary

Black Box models are very attractive from the scientific perspective as they allow they analysis of complex systems without the detailed knowledge of their internal workings. They allow the description of the investigated entity as a number of input and output variables.
The proposed General Black Box Model of HIB builds on existing cognitive theories, introducing an abstract-thinking framework and mathematical description language for HIB related problems.
The key observation of the model is the fact that the user is carrying out his mental processes based not on the objective analysis of his own state, task and environment, but rather on his subjective perceptions. On the other hand this mental process is fully deterministic and perfectly rational.
The analysis of the moderating-factor approaches to IR (e.g. approaches that intend to predict human IR behaviour based on a small number of measurable factors) reveals their numerous limitations such as: • The inability to capture user perceptions • The inability to achieve systematically high predictive accuracy with the number of factors which are omitted in the prediction process • The inability to cope with the high sensitivity of behaviour to input variable change • The inability to cope with variable interactions
The method of problem description introduced by the Black Box Model is going to be used throughout the rest of the thesis to describe the problem of HIB Consistency. It will be used to support the validity of the proposed research framework: the "Integrated Framework for HIB in-situ".

Table 3.1: Summary of chapter 3

4

Achieving personalisation and contextualisation through behavioural consistency

4.1 Introduction

The previous chapter has adopted the Generic Black Box model in order to systematise the discussion about the limitations of the moderating-variable approach to IS&R research. It illustrated the significant complexity of the problem which made the implementation of a generic IR system almost impossible.

This chapter focuses on the concept of environmental and human behavioural-consistency and discusses how this consistency can help to achieve effective IR solutions. The chapter opens by discussing how the concept of consistency along with the principle of behavioural aggregation can simplify the process of analysis and interpretation of human behaviour. The discussion is supported by the geometrical interpretation of the concept using the proposed Black Box model. The model helps to understand how the use of consistency in IR development can help to simplify the moderating-variable space in terms of the dimensionality, as well as the shape of the space itself.

The chapter continues with a review of the cognitive research that successfully takes advantage of human and environmental consistency. It discusses the similarity of the discussed methodologies to the current models of HIB. It also analyses the implications of the principle of behavioural aggregation on the IS&R application design and implementation.

The chapter closes with the discussion of the proposed methodology and its mapping to the general Black Box model. The discussion focuses on the foundations of the idea but also goes through the steps required to develop a successful IR solution. The proposed methodology is the core of the work discussed in this thesis that will further be used to conduct the user studies in order to validate this research.

4.2 Behavioural consistency and the Black Box model

The previous chapter of this thesis discussed in detail the limitations of the moderating-variable approaches such as personalisation or contextualisation. The review carried out in the previous chapter revealed a significant interaction between an entire family of moderating-variables and the users' observed behaviour. This directly affects the predictive validity of the moderating-variable approach. Such high level of interaction between the HIB and the moderating-variables makes the process of predicting human behaviour very vulnerable to the change of those variables. Should any of the moderating-variables (such as the user, context or task) change to any degree, this will have an immediate and difficult to predict effect on human behavioural-responses. In such a case, in order to deliver effective personalized or context-sensitive IR solutions, the research will not only have to understand the entire shape of the human information-related decision-making process, but also to create a model that will be able to cope with its dynamics. Furthermore, the model would have to introduce a user acceptable-estimation error. Using the vocabulary introduced by the Black Box model, it means that the shape of the entire behavioural space has to be understood with an error (Figure 4.2: Error($[\overrightarrow{C_m}, \overrightarrow{P_m}, \overrightarrow{U_m}]$) acceptable by the user. Figure 4.2 b) illustrates the above graphically. The solid line represents the real user behaviour function and the dashed line represents the model used for the implementation. The distance between the solid and the dashed line is the function of error. As a consequence, the overall goal of all of the IS&R, as well as the IR research, is to minimise this error function. The complexity of the problem discussed in the previous chapter, makes the global minimisation of the error function difficult to achieve.

The assumption of environmental and user-behaviour consistency significantly simplifies the above mentioned minimisation problem. If users tend to work in similar contexts, perform similar tasks, or solve similar problems on a daily basis, then in order to deliver effective IR systems it is not necessary to know the entire behavioural function (see Figure 4.2 a), nor to model it in the entire domain (see Figure 4.2 b).

Should there exist even a very limited level of consistency of moderating-variables, then for the purpose of IR system development, the minimisation of the error function can be carried out locally and not globally. This has significant implications which are best described using the Black Box model (see Figure 4.2 c). The existence of consistency means that:

- The moderating-variables which are identified as consistent will not vary in the entire possible range. In such a situation, only a limited set of values can be selected for research and implementation purposes. For example, when investigating the work-related behaviour of software developers, we do not have to understand the impact of user age in the entire available spectrum (i.e. 0 years to 100+) but can focus only on a representable sample. Similarly the task-specific domain expertise does not have to be investigated through the entire spectrum, neither does the task distribution, or distribution across all potential development languages, platforms or technologies have to be investigated. In reality, when focusing on such narrow user and activity samples, a significant amount of variables will not vary at all, or at least only to a marginal, easily controllable extent.
- For each consistent situation, there will exist a number of moderating-factors whose variation will not have an impact on user behaviour, or for which the behaviour variation will be marginal.

$$\forall CPU_{Cons} \in CPU \quad \exists CPU_{Redundant} \in CPU : f(CPU_{Cons}, CPU_{Redundant}) \approx const \tag{4.1}$$

Figure 4.1: Invariability of behaviour across consistent situations

Such variables/components of minor variance can be ignored during the research process significantly reducing the dimensionality of the problem (e.g. through principle component analysis). In practice, the assumption of consistency will allow us to ignore almost an infinite number of variables that do not have an impact on this instance of behaviour.

- The ability to focus only on a very small section of the user interaction-space, allows us to simplify the model used for its interpretation, without having a significant effect on the error. The example illustrated in Figure 4.2 c) shows how the complex model of human responses can be piece-wise linearly approximated. Figure 4.2 c illustrates how the problem complexity can be reduced by an introduction of simple context, task and user-centred heuristics, which are represented by straight lines.

The existence of behavioural consistency does not mean that the developer/researcher will always be able to use the above mentioned simplification techniques. It may also not guarantee an error-free solution. It provides us with a set of tools supporting systematic research and allowing for:

- Reduction of the range of the moderating-factors which are under investigation
- Reduction of the dimensionality of the problem
- Simplification of the model representation

It allows to achieve model simplification without losing control over the error function and provides us with a basis for very accurate error measurement, which is important for the design of IS&R systems and user interfaces.

By designing IR solutions for particular user groups amd performing clearly-defined tasks in well-defined contexts, we limit the variation of contextual / moderating-factors (Figure 4.2 c)) and are capable of creating simple, sometimes even linear models of HIB, and use them to provide the users with a better and more personalized experience. This approach (i.e. the Cognitive Approach) has already been applied with significant success by the biggest Internet service providers such as:

1. Amazon: Designed for all users in a context of intended purchase. The system interprets the act of item purchase as a strong indicator of relevance; and the act of visiting a product site as a weak indicator of relevance. It implements an implicit-feedback recommendation system to provide the user with information about the potentially related articles. Due to its design for multiple users, tasks and contexts it introduces a significant error which is mitigated against by the user interface-design.

These images have been removed

Figure 4.2: Domain Expertise VS Efficiency - a) Efficiency vs Familiarity level (Kelly and Cool, 2002); b) Efficiency vs search expertise and Cognitive style (- Field Dependent Participants, - Field Independent Participants) (Kim, 2001*b*)

2. Gmail: interprets the acts of reading an email; responding to an email; tagging an email; and searching for an email, as indicators of relevance. This information is combined with explicit information about the email importance, provided by the user, and is used to automatically detect emails that can be important to the user. The information is presented in a separate view or can be used during email search up to the discretion of the user. The UI design minimises the negative impact of errors in the estimation process.

3. Last FM: interprets the act of typing in an artist name and then according to the consistent listening of the music provided by the system uses both of these as a source of implicit music-related relevance information. It also combines the information about the music the user listened to, in order to create music profiles which can be shared with other users who have a similar profile.

4. Google: interprets the act of creating a link on the web site as a very strong indicator of contextualised relevance. The link information is used to provide the user with the relevant content. The system minimises the impact of error by providing the user with summary information about the multiple pages the user has retrieved which enables simple reformulation of the query.

The above discussed IR systems simplify the behavioural model of the user by focusing on very specific task and use contexts. It is important to notice that those context-specific implementations are not easily transferable to other, sometimes even very similar contexts.

1. Amazon: The recommendation solution based on previous user purchasing-patterns is not easily transferable, for example, for the purchase of a house or a used car, because the information about the previous purchase-patterns of the user is not available. Even if it was available it makes little sense to recommend houses or cars that have been sold to the user.

2. Gmail: The solution adopted in Gmail is not applicable outside of the email domain, and even within the widely understood email domain it has limited application. In enterprise email, where the circulation of correspondence is very strictly controlled the mechanism for importance annotation has significantly lower value, due to multiple email usage policies.

3. Last FM: Even though the user-profiling approach is transferable to other domains it requires careful consideration and customisation. It is applicable only in the environments where the user is identifiable (not anonymous) and where content-related metadata exists (author, genre, etc.). This makes the solution adopted by Last FM very difficult to adopt in similar applications such as YouTube, where the user interaction is not only frequently anonymous, but also the content metadata is unknown, or even if estimated, carries a significant amount of uncertainty.

4. Google: The relevance-estimation solution adopted by Google is widely recognised to have significant limitations outside of the WWW environment. In the enterprise Intranet context the document authors have significantly different priorities, which have significant impact on the types of documents that are created, as well as their format and the number and types of links that they have on average. Authors perceive links in a different way, and the act of placing them in the document has a different correlation to relevance. As a consequence, the approach which is very effective in the WWW context is not easily transferable to the enterprise search domain.

The above discussed systems demonstrate how consistencies can be taken into account when implementing the IR solutions. However, they do not:

- provide a systematic methodology of achieving the success of implementation in other domains;
- illustrate how to identify behavioural and environmental consistencies;
- provide a guide on how to identify the factors which do not affect behaviour;
- show how to create simplified behaviour models.

The above mentioned implementations were heavily driven by the business needs of the quoted companies and were only created by a significant investment effort which did not follow any particular IS&R model or methodology. They were based on the significant experience of people working within the particular business sectors/user groups, who had significant insight into their customer/user behaviour. They have identified the behavioural-consistencies and took advantage of them by implementing the IR systems. The next section discusses a methodology that provides systematic ways to design personalised and contextualised IR systems. The methodology allows to first identify the business problem and user behavioural consistencies in order to transform them into working and financially-viable IR solutions.

4.3 Existing methodologies

The notion of user behavioural-consistency is a foundation of multiple IS&R theories and models. Multiple HIB models show a significant impact of moderating-factors on the user decision-making process. One of the key modern theories discussing the effects of information focused behavioural moderation is the Holistic Cognitive Theory for IR (Ingwersen and Järvelin, 2007). The theory identifies significant limitations of the laboratory approach which ignores moderating-factors. It proposes a framework that allows the achievement of controllability similar to that of the laboratory approach but taking into account various dimensions of user, task and context. The discussed framework claims to be, and to a significant extent is, an integration of the main IS&R models and research frameworks. As a consequence it is very representative for the purpose of the discussion on consistency and its impact on the IS&R research.

4.3.1 Cognitive Framework

The discussed holistic approach, illustrated in Figure 4.3 b), and described in more detail in chapter 2, is based on the observation that the classical Laboratory Framework of IR is very limited in terms of its ability to cope with real life IR. Real life IR in contrast to the idealistic controllable laboratory model is not an isolated and population-wise uniform process. The authors of the cognitive framework notice the significant importance the moderating-factors have on the users' subjective relevance estimation and build their theory based on the following presumptions (Ingwersen and Järvelin, 2005*a*):

1. "Information processing takes place in the senders and recipients of messages"

2. "Processing takes place at different levels;"

3. "During communication of information any actor is influenced by its past and present experiences (time) and its social, organizational and cultural environment;"

4. "Individual actors influence the environment or domain;"

5. "Information is situational and contextual."

What is even more important for the research presented in this thesis is the fact that the authors of the Cognitive Framework repeat the observations of the cognitive research regarding the interaction between the moderating-variables. They note that the factors affecting HIB are not independent and that there indeed exists interaction between the behavioural variables.

The proposed framework, in contract to the cognitive research, is a constraint to a number of moderating-variables. The framework classifies the key moderating-factors into 9 dimensions/groups. Each of those dimensions is represented as a separate section in Figure 4.3 b). The framework suggests that the selected dimensions could be the key factors responsible for behavioural variance in the domain of IS&R. The framework claims that the 9 dimensions proposed explain a significant proportion of behavioural variance. It suggests that by using the 9 dimensions identified by the cognitive framework, the laboratory approach can be naturally extended towards the user, task and context.

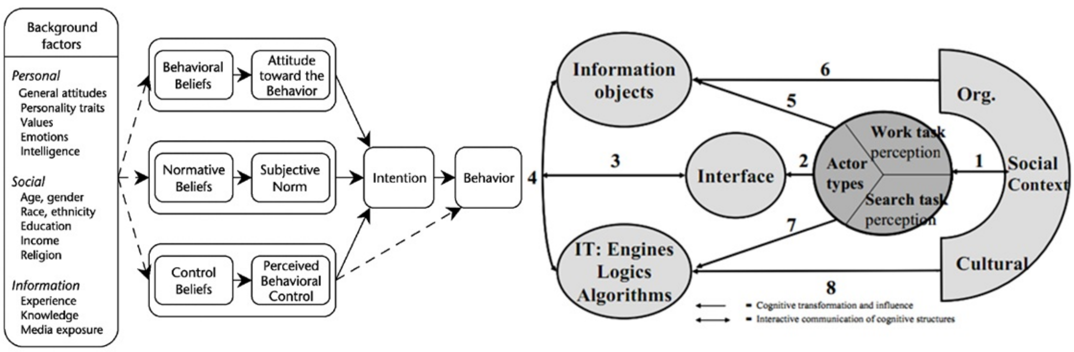


Figure 4.3: Contextualised Theories of Human Behaviour - a) Icek Ajzen's Theory of Planned Behaviour model and the impact of background factors on the model b) Ingversen's and Jrvelin's holistic Cognitive framework for Information Retrieval

The introduction of the 9 dimensions of the cognitive framework promises the ability to "apply a firm hand"on to the context variables. It promises the ability to design laboratory-like experiments that will allow the controlling of the widely understood context.

As discussed in the previous chapter, the cognitive research for decades has tackled the limitations of the moderating-factor approaches which are known for their complexity and very selective nature. And indeed, as easily observe the Cognitive Framework to IR, when generalised, reduced itself naturally to the classical moderating-factor approach. As a consequence it inherits all of its advantages and strengths but also suffers from all of its limitations. The limitations as discussed in the previous chapter relate mainly to:

- Selective nature of moderating-variables
- Difficulty of interpreting the moderating-variable spectrum
- Higher order interactions between moderating-variables
- Number of moderating-variables

Naturally the key limitations were noticed and discussed by the authors of the framework themselves who noted that the biggest criticism of this approach is the fact that (Ingwersen and Järvelin, 2007):

> "there are too many seeking contexts with too many possible combinations of systems and tools: The design and evaluation of IR systems becomes unmanageable."

The above does not mean that the Cognitive Framework is just a very limited, IS&R focused version of the moderating-variable approach, to behavioural research. What makes the Cognitive Framework different, is the emphasis it places on behavioural consistency.

The Cognitive Framework allows the researcher to take advantage of the all the benefits of behavioural consistency because:

- It offers a contextualised approach to the laboratory framework allowing the research to apply a firm hand to reduce the range of the moderating-factors which are under investigation
- By focusing only on the high level variables of the model it reduces the dimensionality and as a consequence complexity of the problem
- The simplification of the model representation does not impose the need for generalisation of research for particular points in the cognitive space to allow for any possible representation

The Cognitive Framework provides a set of tools that can be used when designing experiments targeting particular user groups, consistently carrying out similar tasks in a very consistent (purposefully optimal) manner but it is not straightforward to use. The complexity of the approach is proportional to the complexity of the task of identification of all moderating-variables that are responsible for the variance (or at least its big proportion) for a particular user group and activity. Using the vocabulary of the Black Box model, the key complexity lies in the identification of such a combination of input parameters that would meet the Equation 4.1. The limitations listed above, as well as in the previous chapter, make it very complex to control a sufficient number of variables, in order to both, explain a high proportion of behavioural variance and to implement accurate contextualised IR solutions. Even though very complex, the approach is perfectly viable. The impact of the variation of uncontrolled variables can be significantly decreased (e.g. the application of the Principle of Aggregation (Rushton et al., 1983)) to the UI being designed (e.g. by decreasing the cognitive cost of low-relevance estimation quality or quality variations).

4.3.2 Theory of Planned Behaviour

The Theory of Planned Behaviour (TPB) is an approach (shown in Figure 4.2 a) that explains/predicts human behaviour in specific contexts (Ajzen, 1991, 2005, 2011; Armitage and Conner, 2001). TPB was developed to overcome the limitations of the moderating-factor approach and simplify the behavioural research process (Fishbein and Ajzen, 1975). Similar to the moderating-factor approach, TPB is based on the assumption of human behavioural-consistency. In contrast to the moderating-factor approach, in TPB, human consistency or rationality is a key point that allows behavioural analysis. The theory acknowledges that the human decision-making process is based on the analysis of the environment. In other words, it is based on the analysis of the moderating-factors.

It can be noticed in Figure 4.2 a. that even though moderating-factors are at the foundation of the decision-making process they do not directly drive the behaviour only intentions can do that. Moderating-factors affect human intentions indirectly. This observation, even though very simple, is ground-breaking. Behavioural intentions are derivatives of moderating-factors which are already captured, quantified and processed by the user when actions are formulated. Importantly, when the human cognitive process creates and exposes the behavioural intention to the user, it already takes important moderating-factors into account and therefore their further processing is not required.

In the light of the theory the only other factor, apart from behavioural intentions affecting the actual behaviour directly, is the level of the perceived behavioural control. Indeed in real life almost all of our behavioural decisions are constrained by a variety of factors, both real and imaginary. In real life situations our intentions are moderated by the availability of time, the amount of financial resources, skills and many other factors. The theory states that the stronger the intention and the stronger the perception of control, the higher will be the probability of the behaviour being carried out.

It does not mean though, that there always exists a correlation between intentions, perception of control and the observed behaviour. The correlation will exist only if a number of pre-conditions are met. The following are the key preconditions which significantly affect the predictive validity of the approach (Ajzen, 2011).

1. The key condition relates to the compatibility of intentions, the perception of behavioural control and the investigated behaviour itself. The intentions, perception of control and behaviours are compatible if they are all related to the same behaviour and are observed/carried out in the same context and within a reasonable period of time. To illustrate this with an example based on this research, the intention to copy information from the web related to the current development problem is not compatible with the behaviour of copying the code examples or reusing open-source libraries. The perception of behavioural control is also not

compatible in this case, as different constraints will exist around the reuse of short simple snippets and entire frameworks. Those two information objects will have different intellectual property constraints, complexity and associated cost. Consequently, even though the intentions associated to them can be identical, the different levels of perceived behavioural control will change the decision, of whether to engage in the behaviour of copying. The impact of context on behaviour is not insignificant as well, as the personal opinion of developers towards code or solutions will vary depending on their environment (private/professional).

2. The second condition relates to the stability of intention and the perception of behavioural control in time. The intentions and perceptions of a given behaviour can change in the time interval between their measurement and the observation of the behaviour itself. This condition is the key aspect of this thesis and translates into the requirement of user and environmental consistency.

3. The final pre-condition is concerned with the impact of the perceived behavioural control on the user behaviour. Particularly, the extreme values of the perceived behavioural control would have significant impact on the observed behaviour, regardless of the intentions.

It is also important to notice, that the extent to which the intention and the perception of behavioural control will correlate with the observed behaviour, will vary across multiple tasks, contexts and the behaviours themselves. Usually, with a higher level of perceived behavioural control, the correlation between the behaviour and intention will also increase.

The discussed theory does not only focus on the correlation between intention and perceived behavioural control to the observed behaviour. But importantly, it describes the key factors affecting the intention creation process which are:

- The attitude towards the behaviour - the extent to which user wants to engage in the behaviour

- The subjective-norm/subjectively-perceived pressure to engage or not engage in the behaviour

- The perceived behavioural control - the subjective perception of control the user has over the behaviour

It also discusses why people hold the above discussed attitudes, norms and perceptions of control and goes into much detail of how beliefs shape those perceptions of the world. The theory discusses how our personal beliefs shape our attitudes and norms and how they formulate the basis for our perception of control as well.

The theory discusses the mechanism of creating the behaviour based on our intentions and the perception of behavioural control. It goes further to discuss how our intentions are formulated from both subjective norms and our attitude. Finally, it discusses how the above are created from our beliefs. The understanding of this mechanism is important in the context of this research for multiple reasons as briefly described below:

1. It supports our assumption of temporal user-consistency by placing the rationality of the decision-making process in the centre of the theory.

2. It provides us with a mechanism to infer and investigate the probability of behaviour based on various independent pieces of information especially where information is not directly available.

3. It simplifies and changes our perception of the moderating-variables approach. The Theory of Planned Behaviour also takes advantage of moderating-variables. Importantly, the moderating-factors affect the observed behaviour only to the extent that they affect the limited number of beliefs that are shaping our intents. They only affect the behaviour to the extent that they actually shape our beliefs.

The above reasons have significant implications from the perspective of the Black Box Model analysis of the problem. The assumptions of the Theory of Planned Behaviour align with the assumptions of the Black Box Model. TPB allows identifying a limited and finite set of beliefs that are responsible for a significant proportion of behavioural variance. As a consequence the identified beliefs allow the identification of a finite number of moderating-variables that explain their variance. When a set of beliefs (moderating behaviour) and moderating-factors are identified then both the range and the dimensionality of the moderating-factor space is significantly reduced. Additionally the Theory does not impose any limitations on the model used for describing the behaviour under the varying moderating-factor context. In practice though, the majority of models used along with the TPB are linear thus simplifying the problem.

The TPB's predictive validity has its limitations as it rarely achieves reliability ranging from 0.75 to 0.80 (Pearson's p) (Ajzen, 2011; Fishbein and Ajzen, 2009). It also has very well understood constraints of use, as discussed above.

The TPB is a very mature approach to the problem of behaviour prediction and control and as such has been subject to criticism primarily from the supporters of the moderating-factor/Black Box Model approach. One of the strongest criticism has been raised by Fazio (1990) who states that, as the name indicates the TPB focuses primarily on the planned/intentional behaviour (refer also to Aarts and Dijksterhuis (2000); Knussen et al. (2004); Ouellette and Wood (1998); Wood et al. (2002)) He claims that the TPB ignores the effect of automatic behaviours and processes. They believe that the observation of habitual behaviour that is the behaviour recurring frequently and consistently across the same contextual settings that can be triggered by environmental queues (Aarts and Dijksterhuis, 2000) can explain the same amount if not more variance than the TPB. They believe that the habitual behaviour is triggered without intention, which is the foundation of their criticism as TPB does not explain such behaviours. This point of view is challenged successfully by Ajzen (2002) who points out that the habit is a result of systematic repetition of deliberate action and can not exists without a strong intention even if at a later stage the execution of the task is supported by some automated mental processes. Ajzen provides a comprehensive body of evidence that the residual impact of past behaviour vanishes when intentions are strong (i.e. when we do not want to engage in the behaviour any more e.g. because of the contextual/task change).

Similarly TPB is also criticised for its inability to cope with emotional responses (Arvola et al., 2008; Bae, 2008; Batra and Ahtola, 1991; Edwards, 1990; French et al., 2005; Trafimow and Sheeran, 1998). This and the previous criticism can be further generalised and discussed in terms of the "Sufficiency of the Theory of Planned Behaviour" (Ajzen, 1991). There exists a substantial body of research aiming to extend the TPB with tens of other moderating-factors, presumably in certain situations increasing its predictive validity. In this light it is important to make the following two comments citing Ajzen (1991):

- The 'theory of planned behaviour is, in principle, open to the inclusion of additional predictors if it can be shown that they capture a significant proportion of the variance in intention or behaviour after the theory's current variables have been taken into account'

- 'it is at the level of beliefs that we can learn about the unique factors that induce one person to engage in the behavior of interest and to prompt another to follow a different course of action'

The TPB has also a more fundamental/philosophical limitation which is founded in a way we define action and more precisely the factors triggering and controlling it. As discussed by Suchman (2007) there are two alternative views of action:

- One that explains action through the underlying plans. As one can clearly see the TPB falls into this category along with the wider behavioural sciences.

- An alternative view proposed by Suchman (2007) suggests that "although the course of action can always be projected or reconstructed in terms of prior intentions and typical situations, the prescriptive significance of intentions for situated action is inherently vague. The coherence of situated action is tied in essential ways not to individual predispositions or conventional rules but to local interactions contingent on the actor's particular circumstances."

The second view action recommends and ethnomethodological approach that shifts the focus on analysis to: action, the particular and concrete circumstances in which it occurred and the process in which people interpret the context in order to carry out action. "Rather than build a theory of action out of a theory of plans, the aim is to investigate how people produce and find evidence for plans in the course of situated action." Suchman (2007).

The above does not mean that the two proposed approaches are conflicting or mutually exclusive. On the contrary in the author's opinion they can be used together to gain a detailed understanding of an activity in a certain situational and social context (through ethnomethodology) and provide very solid foundation for generalization through application of TPB (or similar) to the identified behaviours (reported levels of correlation across multiple studies are very high and vary between R=0.59 and R=0.66 (Ajzen, 2011; Fishbein and Ajzen, 2009)). The proposed methodology is a form of such integrated approach (for an explanation of this, refer to the last section of this chapter).

The above is a very appealing approach as the TPB has historically proved effective. A meta-analysis carried out by Armitage and Conner (2001) reviewed more than 180 independent studies taking advantage of TPB to predict user self-reported and observed behaviour. The results found $R = .52(R^2 = .27)$ for the multiple correlation of intention and perceived behavioural control with behaviour. Older meta-analyses are consistent with those findings with R ranging between .46 and .58 (Notani, 1998; Rivis and Sheeran, 2003; Schulze and Wittmann, 2003). The methodology of TPB is successfully used in a variety of domains such as prediction and regulation of health-related behaviours. TPB is one of the key models of Social Marketing for the prediction of consumer behaviour, or prediction of pro-environmental behaviours.

Based on the above, it is considered that the TPB will benefit IS&R research and IR implementation provided that the behaviours under investigation, population and the proposed system constraints are defined prior to the application of the TPB. The TPB does not support the steps needed to identify those behaviours due to the fact that it is a behavioural control theory.

The next section will discuss the proposed "Integrated Framework for HIB in-situ" that takes advantage of the two discussed frameworks/theories to provide this research and a broader research community a framework for effective IS&R experimentation and IR design.

4.4 Proposed methodology: "Integrated Framework for HIB in-situ"

This section proposes an integrated methodology based on two existing well known methodologies as discussed in the previous sections. The reviewed methodologies have significant advantages but also their own limitations as listed in Table 4.1.

By integrating the two discussed methodologies one can benefit by taking advantage of the strengths of the two methodologies and minimise their limitations. The proposed approach has following advantages:

	Theory of Planned Behaviour	Holistic Cognitive Theory for IR
Advantages	Very high predictive validity. Very simple research methodology. Widely available set of research tools and application examples. Very strong psychological foundations. Wide applicability across multiple research problems/domains. Well established history of research.	Ease of applications in the IS&R context. Strong HIB and IS&R foundations. Well embedded into IR system development process. Actor and behaviour focused.
Limitations	Limited applicability for the IR development process. Designed for investigation of very well-defined instances of human behaviour.	Limited history of prior applications. Limited evidence of predictive validity. Moderating-variable approach (suffering from all of the approach limitations).

Table 4.1: Advantages and limitations of the Holistic Cognitive Approach and Theory of Planned Behaviour

- It takes advantage of the Cognitive Holistic Framework in order to simplify the process of identification of:
  - User groups involved in the retrieval process, the key contexts in which IR takes place and tasks which are being carried out
  - Interface and information object preferences as well as the behaviours following retrieval
- It takes advantage of the high predictive validity of the Theory of Planned Behaviour in order to obtain a clear understanding of the moderating-factors explaining a big proportion of variance. It also provides us with tools that allow the understanding of the amount of variance that is not explained. This drives the key design decisions.
- Overall it is simple and clear and when supported by flexible research tools and examples, has a significant potential for use outside of the academic domain leading to a significantly decreased cost of effective IR development.

Both models work together and focus on separate aspects of the IR development problem being the investigation/prediction of behaviour and the embedding of the knowledge of user behaviour into the IR system. Importantly the proposed methodology provides a systematic set of steps, that can be followed to deliver an effective IR solution decreasing significantly the research overhead of the IR implementation.

Similar to the Theory of Planned Behaviour and Holistic Cognitive Theory for IR, the 'Integrated Framework for HIB in-situ' is based on the assumption of Human Behavioural Consistency, hence it was illustrated as a foundation of the approach in Figure 4.4.

This image has been removed

Figure 4.4: Domain Expertise VS Efficiency - a) Efficiency vs Familiarity level (Kelly and Cool, 2002); b) Efficiency vs search expertise and Cognitive style (- Field Dependent Participants, - Field Independent Participants) (Kim, 2001*b*)

The proposed framework focuses on the "behaviour" which is the common point of the two underlying theories. In the proposed methodology it becomes a bridge mapping the two methodologies on to each other. Its consistency is the presumption of the methodology and the focal point of this research. The proposed methodology based on the framework is illustrated in the bottom section of Figure 4.4 as an arrow representing the order of activities. The methodology is composed of the following steps, taking advantage of both of the underlying frameworks.

1. In Step 1, the Cognitive Heuristic Framework is used to identify the IS&R actor/actors and the work and search tasks for which the further work will be carried out. This goal of this stage is to define the exact scope of the research and its importance should not be underestimated. At this stage the drive/business case for the research / the IS&R problem is identified along with the measurable metrics of IS&R improvement/commercial success.

2. The identification of actors and tasks is then followed by the detailed work analysis. The work analysis should lead to a detailed and documented understanding of the processes involved in IS&R, performed by the selected group, in the selected context. It should result in the identification of currently used information objects and interfaces and their limitations. Also this stage should provide the research with a detailed understanding of the purpose of the retrieved information and how this information further feeds into the work process. Special attention should be placed on the revision of behaviours that follow the successful retrieval as those can be further used for the development of various "pseudo-relevance-feedback" prototypes.

3. The work analysis ends with the identification of interfaces, information objects and/or behaviours that can be used in the new IR systems based on contextualisation or personalisation. This is a starting point for the next phase of research when the repeatability and consistency of the identified behaviours will be assessed.

4. The subjective preference of interfaces and information objects, along with the necessary behaviour, is then further analysed using the approach consistent with the TPB, in order to identify the key beliefs impacting the behaviour. This step is essential for the identification of stable behaviours as well as the identification of the level of variability of the behaviour. Additionally, key beliefs and factors

that will affect the extent of variability are taken into account. During this step the researcher has to understand which behaviours are consistent, whether the proposed implementation can cope and to what extent, given the levels of consistency observed.

5. The identified beliefs leading to the target behaviours can be used in the next step by the identification of the key moderating-factors which affect their behaviour. The understanding of moderating-factors and the extent to which they can be captured measured and controlled, leads naturally to the design decisions and the introduction of measures of control, or the mitigation of their variability. It is important to notice that the variability of behaviour, combined with the inability to control the moderating-factors affecting them, does not necessarily discriminate the behaviour. Various methods can be used to either decrease the variability (i.e. the principle of aggregation), or to decrease the impact on the user (e.g. by decreasing the cognitive load associated to review of the incorrect information).

The analysis of the proposed methodology from the perspective of the Black Box model proves to be very consistent as well. Every step of the proposed methodology translates directly into dimensionality reduction, variable-range reduction or model simplification activities. The methodological steps of the framework can be described using the Black Box model vocabulary in the following way:

1. Actor and problem identification: This step is equivalent to the reduction of the variation of the key dimensions of the behavioural space.

2. Work analysis: The identification of a group of behaviours as a result of this activity allows for the short listing of functions that will be modelled using the Black Box model.

3. Identification of behaviours: The selection of a single behaviour: This step is equivalent to selecting only one of the behavioural functions to be modelled in further research.

4. Behavioural analysis: This is consistent with the identification of groups of dimensions that have a key impact on the selected behaviour consistency. This step

reduces the size and dimensionality of the behavioural space under investigation even further. Moreover, during this step the nature of the relationship between the dimensions of the model and the observed behaviour is revealed.

5. Identification of moderating-factors: The identification of key dimensions that will have impact on the selected behaviour consistency and the investigation of the relationship between those dimensions and the behaviour function itself.

The next chapter of this thesis will review the fundamental assumption of the proposed methodology being the assumption of behavioural consistency and will illustrate how the proposed consistency focused methodology can be used during the research and development process. This will be achieved by the application of the discussed methodology to the problem of IR, in the context of software development.

Chapter 4: Summary

The analysis of the problem of HIB Prediction using the General Black Box Model reveals that its complexity can be significantly reduced if the prediction is carried out locally and not globally in the problem space.
The analysis of HIB for most frequently reoccurring/consistent areas of the Black Box Model space (reoccurring tasks/contexts/emotional states) allows the: • Reduction of the dimensionality of the problem • Reduction of the variability of the remaining variables • Decreasing of the sensitivity of behaviour to variable change • Reduction of the complexity of the problem representation (e.g. through liberalisation)
Taking advantage of user consistency can lead to the significant simplification of the problem of HIB prediction and the development of highly effective personalisation/contextualisation systems. All of the successful personalisation/contextualisation systems (some of which are analysed in this chapters) target their implementation to well-defined and highly repeatable/consistent behaviours.
Consistency of human behaviour is also a focal point of numerous Cognitive and IS&R theories/ research methodologies such as the Cognitive Holistic Framework for IR or the Theory of Planned Behaviour.
The discussed frameworks have numerous limitations that make them unsuitable for rapid development of personalised/contextualised IR systems. • The Cognitive Holistic Framework to HIB is a highly consistency focused moderating-factor approach and as a consequence suffers from all of the moderating-factor related limitations as discussed in previous chapters. • The Theory of Planned Behaviour is an abstract cognitive theory and as a consequence is not easily applicable to IS&R.
The proposed "Integrated Framework for HIB in-situ" integrates the above consistency focused approaches. It uses the two distinct theories in order to address their mutual limitations and suggests an approach that not only allows for achieving high predictive validity but is also easily applicable in the IS&R domain.

Table 4.2: Summary of chapter 4

5

Software developers' behavioural consistency

5.1 Introduction

The previous chapter discussed the importance of behavioural consistency for the overall HIB research but also for the process of design and implementation of real life IR systems. It focused on the most representative behavioural change theories from psychology and HIB. It also discussed how they use the notion of consistency in order to describe and predict human behaviour. The review carried out in the previous chapter illustrated, that there exists psychological research discussing how to achieve considerable predictive validity, when predicting human behaviour. However the existing research is not focused on the design and delivery of IR systems. As a consequence, the application of theories such as Theory of Planned Behaviour to commercial IR problems and their integration into the software-design and delivery process, is not a trivial task. On the other hand the HIB theories, such as the Holistic Cognitive Framework to IR, even though designed with the IR problem in mind, lack the maturity of the psychological research and its predictive validity.

The previous chapter pointed out the importance of the assumption of consistency for behavioural research but also identified the significant methodological and theoretical limitations of the existing frameworks. Therefore the previous chapter proposed a new methodology, "Integrated Framework for HIB in-situ". The proposed methodology integrates the strengths of the Theory of Planned Behaviour and the Holistic Cognitive Framework to IR. The proposed methodology helps to simplify the commercial IR delivery process and overcomes some of the limitations discussed previously.

This chapter discusses the application of the proposed methodology by illustrating how it can be applied to the real life information problem. In this case the methodology is being applied to the work-related information problems of software engineers.

The chapter is composed of two key parts as represented by the bottom and top sections of Figure 4.4 shown in the previous chapter:

- It opens with the discussion around the way the key assumption of the research (human consistency) will be verified. It discusses the way the overall human information-related behavioural consistency will be assessed (top of the Figure 4.4). As discussed, the assumption of human consistency is fundamental for the overall behavioural research. The validity of the assumption of consistency in the context of HIB, is being verified at the beginning as a first activity of this research, by User Study 1.
- The following sections of this chapter go through the steps of the proposed "Integrated Framework for HIB in-situ". The discussion focuses on how particular activities of the proposed methodology are being applied in the context of this research (bottom of the Figure 4.4). The chapter discusses how steps such as the identification of actor and task groups, the further investigation of behaviours and their variability through User Studies 2 and 3, are designed and carried out.

5.2 Overall HIB consistency

As discussed in previous chapters, the consistency of Human Information Behaviour is a necessary condition for any systematic attempts, which are aimed at explaining and predicting HIB. The existence of the causal relation between how humans perceive the world around them and how they respond to it, is essential for any of the scientific methods to make sense at all.

The goal of the first activity of this research is to review the assumption of human information-related consistency. User Study 1, described in this section, attempts to investigate to what extent daily HIB is repeatable and therefore to what extent there exists a potential for the development of task specific IR solutions. The User Study

1 was composed of three stages: an opening interview, diary study and the closing interview.

The user study was performed on a group of 12 participants for a duration of two months. The group consisted of 6 women and 6 men. Five participants of the research were postgraduate students of Coventry University, three participants were undergraduate students, and five participants were working in Coventry and the surrounding areas.

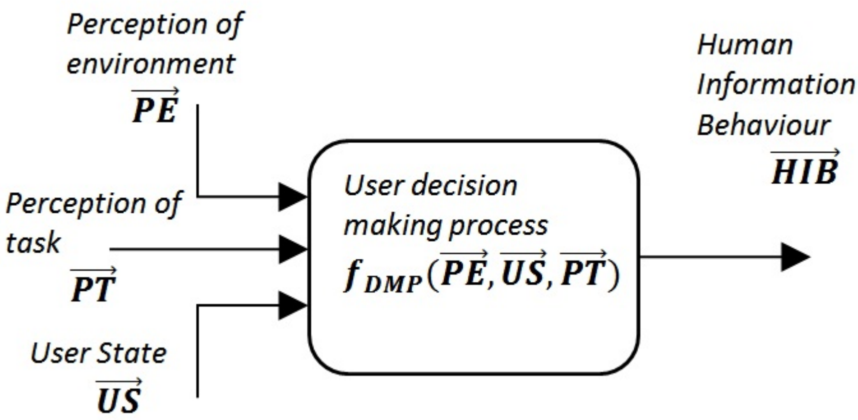


Figure 5.1: Black Box Model - The General Black Box Model of HIB

User Study 1 focuses on the key elements of the Black Box Model of HIB (Figure 5.1) and the extent they are perceived by users as repeatable/consistent/habitual:

- The study focused primarily on the perceived environmental-consistency of HIB. One of the key goals of the opening interview was to establish the existence or non-existence of the geographical HIB distribution. The interview focused also on a number of key contextual factors (moderators) that affect user information interaction as well as its distribution and variation. This part of the uses study focused significantly on the perception of environment ($\overrightarrow{PE}$ vector - Figure 5.1) aiming to prove that the geographical location of HIB not only has a non-uniform probability density function but also it is a sum of normal distributions supporting the assumption of environmental-consistency.

- The study also focussed on the user reported work task distribution both in time and space, concentrating therefore on the second vector of the Black Box Model:

$\overrightarrow{PT}$ (Figure 5.1). The information about user work distribution was collected in a twofold manner, both directly during the interview, as well as through the analysis of the diary entries. The diary entries were used to sanitise the verbal information provided by the users and make sure that the non-key activities (in terms of importance) were not omitted by the users, who tended to focus on their critical tasks.

- Thirdly, the study focused on the distribution of user key states ($\overrightarrow{US}$ - Figure 5.1), and their correlation to the other input and output variables of the Black Box Model. The key focus was on the analysis of the variation of the (US) vector across various points observed in the $\overrightarrow{PT}$, $\overrightarrow{PE}$ space. The analysis focused on the variation of user states across highly repeatable tasks and environments. The variation was hypothesised to be very low which supported the hypothesis of human consistency.
- Finally the study investigated the consistency of user responses, especially to the problems, environments and states self-reported by users as being reoccurring /consistent. This was achieved though the analysis of the user-perception of the problem, its consistency and the use of search tools/ strategies applied as well as the search keyword analysis.

The design of the diaries allowed for the estimation of the magnitude of the observed regularities and their correlation with the selected moderating/contextual factors. Also it allowed the assessment of the regularities in terms of distribution in time and space and across different contexts of interaction. Finally User Study 1 was compared with the general psychological research so that the analysis of the underlying problem-solving and decision-making mechanisms could be carried out and further generalization of the results could be achieved. The design of the study was inspired by the work carried out by Bargh (2007) on the automaticity of higher mental processes, Ajzen's (2002) view of the habituation and automaticity process, as well as Wood's work on habituation (Ouellette and Wood, 1998; Wood et al., 2002). This thesis is not devoted to the automaticity of human behaviour or the process of habit development. However, the results are comparable with the results obtained by Wood et al. (2002) which discuss the extent of human everyday consistency.

5.3 Actor identification

User Study 1, discussed in the previous section, supported the assumption of HIB consistency but also revealed significant variations of the usage of IR systems and the levels of consistency across various user groups. The observations made the application of the proposed methodology ("Integrated Framework for HIB in-situ") viable, but also stressed the significant importance of the careful selection of actors and problems involved in the study, due to the observed variations in consistency. This section discusses the process used for actor identification and supports the selection made, with appropriate evidence from the literature.

The initial identification of the actor and problem group was based on the simple observation of the current employee work patterns, of the sponsor of the discussed research. The day to day practice illustrated clearly, that the people involved in software design and delivery are highly dependent on access to the on-line information resources and that the frequency of the use of those resources is very high. The analysis of the feedback from the teams responsible for code development suggested that the process of finding the information is not always straightforward. Finding information can take a significant proportion of their overall work time and consequently it generates a significant cost.

The literature review supports the observation that software engineering and development on every stage of the software life-cycle is a very cognitively challenging task (Detienne and Bott, 2001; Hoc, 1990; von Mayrhauser, 1995; Weinberg, 1998; Xu and Rajlich, 2004). It requires from the developer a variety of technical, problem-solving and business skills. On the one hand, the process of software design and implementation requires significant creativity and innovation, but on the other hand it can be very tedious, requiring from the developer repeatable activities and a significant amount of manual, often non innovative work. Both the complex and trivial activities are usually executed under significant pressure, to optimise the developer performance, in terms of time and quality of code (Freund et al., 2005; Ye et al., 2007), which directly translates into profit generated by the software development team and the software user.

In late 60's this relationship was identified and new groups of research appeared focusing on the Psychology of Computer Programming (Weinberg, 1998), Information Science and widely understood HCI for software development. The early research focused mainly on the investigation of the cognitive and learning processes involved in the development of code with the special emphasis on how to use this knowledge to develop new, easier to use, programming languages and tools, supporting the development of code. The introduction of computer networks and especially affordable Internet and Intranet significantly changed the way software engineers solve their work-related problems. Due to significant social pressure and availability of online resources and communication channels, more and more information problems in the development life-cycle are resolved exclusively through the use of on-line information and code repositories. The changes in the software development practices and availability of online technologies also triggered this shift. The focus of the research community changed both from Psychology of Programming and Information Science perspectives. Current research focuses significantly on the usage of online resources and the IR process.

The more specialised literature discussing HIB within the Software Engineering and also in a wider Engineering domain is also very mature. This provides very solid support for the research focusing on this user group. The next subsections will focus on the more detailed literature review of engineers' and then software engineers' work practice and the importance of IR throughout the process. The review will allow to understand further the importance of IS&R utilities and provide with knowledge about the frequency with which we can expect to observe seeking behaviours and the estimated extent of their consistency.

5.3.1 The role of IR in Engineers' and Scientists' work practice

The information behaviour of engineers (also including engineering students and researchers) is one of the most investigated topics within the HIB domain. In his 1984 review of the HIB landscape, Wilson et al. (1984) refers to the literature of a time as primarily the "study of scientists' information-seeking behaviour". Because of the availability of engineers and scientists within academia and business focused on IS&R research (e.g. Miscrosoft, Google, Yahoo, Facebook, Amazon, etc.) there exists a wealth of literature discussing the information needs and habits of those two user groups. Table 5.1 provides an overview of the most representative studies on Engineers' and Scientists' information behaviour.

Even though the information-needs as well as goals of Scientists and Engineers are similar, the research has identified differences in the way they acquire and use infor-

User Group	Key Representative Studies
Scientists	Bichteler and Ward (1989); Brown (1999); Case (1986); Ellis et al. (1993); Flaxbart (2001); Hallmark (2001); Murphy (2003); Palmer (1991); Saracevic (1992)
Engineers	Bruce et al. (2003); Fidel and Green (2004); Freund et al. (2005); Gerstberger and Allen (1968); Gralewska-Vickery (1976); Hertzum (2002); Hertzum and Pejtersen (2000*b*); Holland and Powell (1995); Leckie (2005); Pinelli (1991); Tenopir and King (2004); Yitzhaki and Hammershlag (2004)

Table 5.1: Summary of the HIB literature focusing on Scientists' and Engineers' behaviour

mation. As discussed by De Solla Price (1965), the groups differ primarily because "the scientist wants to write but not read, and the technologist wants to read but not write". This has direct implications to the way those two groups search for information. Research carried out by Yitzhaki and Hammershlag (2004) investigated the HIB differences between engineers and scientists in terms of what information resources were used and what factors decided about the information relevance.

Work carried out by Leckie et al. (1996) was one of the most widely recognised within the HIB and IS&R community. The proposed model of professionals' work-related HIB was based on a extensive survey of work practices of scientists, engineers and several other groups of professionals. The work carried out by Leckie (2005) acknowledged the importance of IS&R among the investigated groups and stressed the critical nature of utilities and processes supporting this activity.

The model illustrated in figure 5.2 is a moderating-factor approach to the problem of HIB. The proposed model identified six dimensions of context or moderating-variables that had primary impact on IS&R in work-related situations, which are:

- The work role and task at hand which are inevitably interconnected and together influence a definition of the task related information needs.

- The characteristics of the information need. This element is consistent with the user-centred approaches discussed in the previous chapters of this thesis. According to Leckie it is the key driving force for the overall HIB.

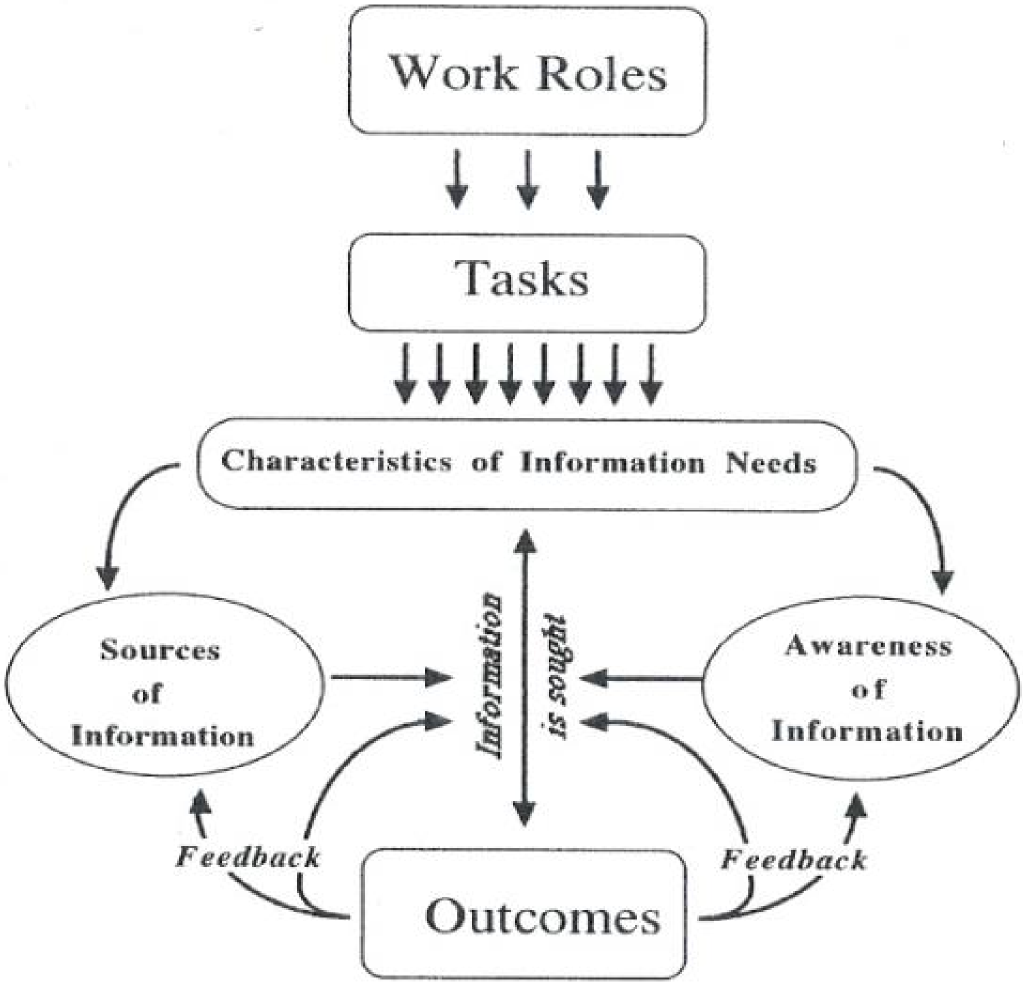


Figure 5.2: Leckie's Model - Leckie's et. al models of information seeking behaviour of professionals (Leckie, 2005; Leckie et al., 1996)

- The awareness of information and information sources. What is recognised by Leckie and what is characteristic of the professional and work-related IR which is not always Internet based, is the fact that awareness of the information plays a very significant role. Awareness of the information along with the sources of information is one of the key mechanisms mediating the process of IS&R and leading to the selection of search outcomes.

- The outcomes and their relation to the above through the process of seeking and feedback.

One of the most consistent observations of the literature investigating engineers' information-behaviour is the fact that engineers prefer processed information over detailed documentation (Case, 1986; Holland and Powell, 1995; Tenopir and King, 2004). The empirical studies illustrated that engineers relied on their own knowledge as well as the knowledge of their collogues (either directly or through the use of blogs or discussion forums) more frequently than on technical documentation. One of the primary reasons fur such behaviour was the problem-driven nature of their work. Engineers often worked under strict time constrains and were heavily task-driven. From their perspective, full understanding of the problem was not as relevant as problem resolution. Engineers would attempt to minimise the time and effort required to gain the additional knowledge. They would use the sources that they were familiar with and that they believed were most accessible. They would reinforce the use of those selected source over time (Hertzum and Pejtersen, 2000*a*).

There also existed vast body of research illustrating the importance of interpersonal communication among engineers. Holland and Powell (1995) have conducted a user study on a group of 60 engineers asking them to rate their information resources in terms of their importance. "Word of moth" was the most highly rated of the listed sources.

The study carried out by Fidel and Green (2004) investigated the process of information selection by engineers. One of the primary findings of the study was the fact that engineers were heavily focused on saving time during IR and selection of documents. Therefore they very frequently preferred familiar information sources even if they provided suboptimal information.

The following section will illustrate that the above observations are also true for software engineers. Software engineers deliberately avoid using the technical literature if they can find highly processed information authored by their collogues and reuse existing solutions (employing the principle of Least Effort). It will also demonstrate the importance of effective IR and the implications it has on their overall performance.

5.3.2 The role of IR in Software Engineers' work practice

The literature review carried out for this research supports the initial findings and consultations made in the sponsor's work environment. Because of the overall complexity of the problem discussed above, even before the widespread use of the Internet, searching for information was one of the most common and also one of the most time consuming activities, carried out by software developers (Curtis et al., 1988; King and Griffiths, 1991; Liu et al., 2005; Seaman, 2002; Singer et al., 1997). The figures quoted in the literature vary significantly across different experimental groups reporting that even 60% (Liu et al., 2005) - 80% (King and Griffiths, 1991) of software developers' time could be consumed by information retrieval. Most often, frequencies between 20% and 30% are reported in the literature.

Paper	From	To	Research method
Liu et al.	60%	60%	Questionnaire supported interview carried out on 9 participants
Brandt et al.	14%	41%	Observation of 20 students participating in a case study
Singer et al.	24%	45%	Questionnaire on 13 participants, shadowing a single participant for 6 months
King & Griffiths	20%	80%	
Freund et al.	20%	30%	Interviews with 14 software engineers / consultants
LaToza et al.	5%	20%	Survey carried out on 187 Microsoft corporation software developers
Perry et al.	16%	16%	A yearlong diary study carried out on 13 participants
Tenopir & King; King et al.	40%	66%	Very wide literature review specific to information use of engineers

Table 5.2: Average amount of time software developers spend on information retrieval / communication

An example of a typical work task distribution of software engineers was reported by Singer et al. (1997) and is illustrated in Figure 5.3. The graph clearly illustrates the importance of information, its re-use and retrieval, in the process of software maintenance and development. The graph clearly shows that even before the widespread of Internet IR was not only the most frequent activity it was also carried out on a day to day basis.

These graphs have been removed

Figure 5.3: Key Software Engineering Tasks - The graph on the left illustrates the percentage of days (out of 14) during which the reported behaviour occurred at least once. The graph on the right illustrates the frequency of particular activities (Singer et al., 1997)).

The widespread availability of digital information did not significantly impact the proportion of time developers spent on search (Brandt et al., 2009; Freund et al., 2005). On the contrary, the increasing complexity of development problems; the technologies used to deliver code; the increasing availability of open source tools/components and easy access to the information resources, further increased the amount of time developers spent on searching for information and engaging in opportunistic programming (Brandt et al., 2009; Hartmann et al., 2008). The research indicates that the proportion of time spent on searching still varies between 20%-30% and would reach much higher levels if it were not for the commercial/management related constraints (Freund et al., 2005).

The role of IR and its distribution in time and across tasks was analysed by Brandt et al. (2009). Figure 5.4 illustrates the results of their work and shows an exemplary distribution of web usage during the software development activities. The results stress the importance of IR activities in the development life-cycle but also reveal the individual nature of IR and information re-use. The amount of time individuals spent on the IR varies between 7.2 and 68.8 minutes. It is important to notice that the Internet was used by all of the participants and all of them perceived it as a "key resource".

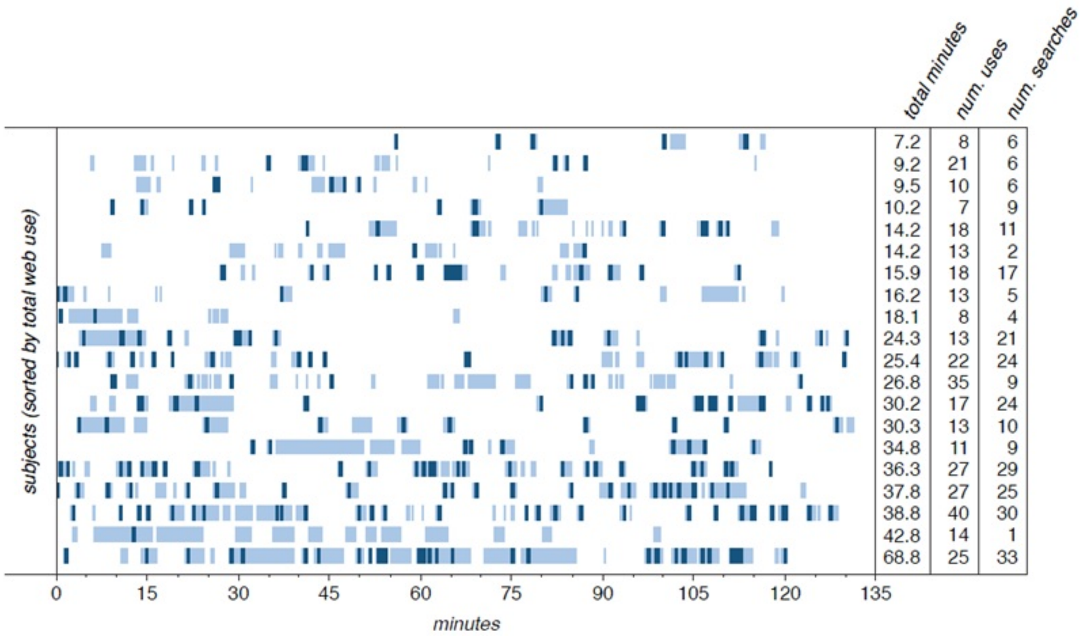


Figure 5.4: Time distribution - Web resources usage during software development task (Brandt et al., 2009). The dark bars represent the time spent on the searching activity; the light bars represent the time spent on assessing the content retrieved by the search process.

The literature review clearly illustrated that not only the initial observations carried out in the sponsor's environment were correct (indeed, IR is one of the core engineering activities), but also showed that the fact that software developers stop the search process is not necessarily related to the success in IR. The fact that software developers spend 20%-30% of their time on IR does not mean that they are able to solve all of their information problems in this time frame (Freund et al., 2005). The amount of time is limited because of business constraints, deadlines and delivery goals which prevents developers from continuing their search activities. This has a negative impact on the code and software architecture quality.

The review illustrated that there exists a significant potential (also commercial) for new and more agile IR solutions, focusing on widely understood software engineering.

The review shows that the group of software engineers is a good candidate for further behavioural analysis due to the frequency of interaction.

5.4 Task identification

According to Ajzen (2002), "the frequency with which behaviour has been performed in the past is found to account for the variance in later behaviour, independent of intentions". This is especially valid if the context in which the behaviour is being executed, along with the human attitude towards the behaviour, is stable (Ajzen, 2002; Ingwersen and Järvelin, 2005*b*; Jrvelin and Ingwersen, 2004; Wood et al., 2002). Obviously, there exist an almost infinite number of moderating-factors, that account for context and user personality variation and they all have a dramatic impact on the decision-making process (Aarts and Dijksterhuis, 2000; Ouellette and Wood, 1998). If the variation of those factors is controlled, for example by the imposed task and the social settings, then a successful prediction of behavioural trends can be achieved with significant accuracy (Ajzen, 2002).

This section focuses therefore on the identification of tasks in which the software engineers are engaged. It shortlists the key tasks and actor groups which have the highest potential for the further analysis of IR development. This exercise is carried out in order to understand the wider context of their interaction and identify, through work analysis, a number of behaviours that will be further investigated in User Studies 2 and 3.

The analysis carried out in this section aims at reviewing the task distribution of Software Engineers and selection of those activities and engineer groups (e.g. testers, analysts, developers, administrators, etc.) for which further analysis and optimisation will be carried out. The selected tasks should have the following properties:

- They should be frequently observable
- They should have non-trivial complexity
- They should naturally transform into search tasks, also of high frequency and non-trivial complexity

In order to achieve the above mentioned goal the literature review was carried out aiming to identify various Software Engineer task and activity classifications. One of the most comprehensive investigations of Software Engineers' task distribution in context of IR was carried out by Freund et al. (2005). The research provided an overview of the changing information needs and strategies used by software engineers to find the required information. The analysis was performed taking into account contextual factors of developers' interaction such as work task category, stage of the project, information goals and many others. What is important in the context of this review, is the fact that the research identified a number of repeatable classes of activities/tasks/archetypes that were characteristic of software development and discussed the information needs for each of them. The results stressed the fact that the Software Engineers' information behaviour was highly shaped by the work context in which the information need was created. The reported work suggested the existence of a causal relationship between the work task, the context in which it is performed and the information behaviour/strategies engaged in searching. This seemed to be supported by the previous literature investigating the professionals' information behaviour (Freund et al., 2005; Jrvelin and Ingwersen, 2004) and the literature discussing the impact of task on information behaviour (Kim and Allen, 2002). Figure 5.5 based on Freund et al. (2005), summarizes the Software Engineers' activities and highlights several dimensions across which they can be classified. The results refer to the wider group of Software Engineers and a set of tasks which are not exclusively focused on code delivery.

The distinction of roles and responsibilities identified by Freund et al. is also widely recognised outside of the IR research domain. The observation that Software Engineers are becoming highly specialised was made also in the Psychology of Programming. The review carried out by Freund and many others pointed out that Software Engineering involves a vast range of various activities which are not necessarily related to the delivery of code.

The further literature review illustrated that the role of a Software Engineer was not necessarily assigned to him permanently but can change along with the life-cycle of the project (LaToza et al., 2006). To understand the task distribution as well as tools and most common practices of Software Developers, researchers LaToza et al. (2006) and Kim (2006) conducted two surveys and eleven interviews on 187 software developers from Microsoft Corporation. The experiment was focused on employees engaged in the

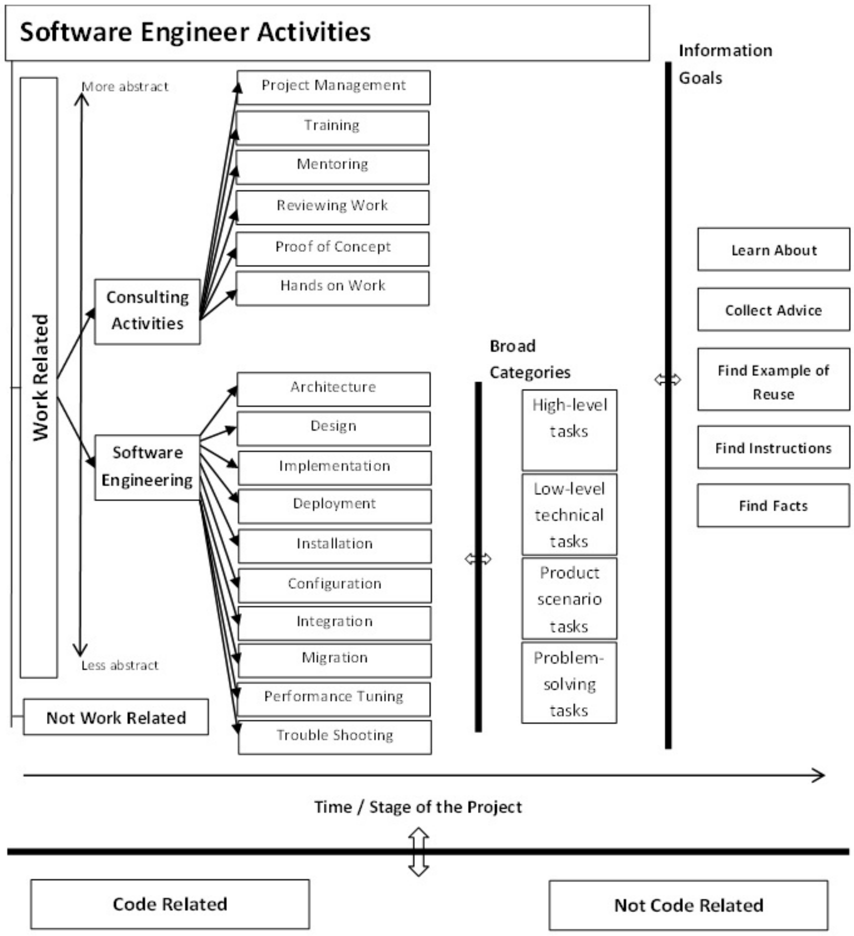


Figure 5.5: Classification of Software Engineer Work - Software Engineer Activities distribution based on Freund et al. (2005)

code delivery process, therefore the results were less abstract than those obtained by Freund et al. Figure 5.6 (a) illustrates this distribution and the percentage of time consumed by different tasks.

This image has been removed

Figure 5.6: Software Developers' Activities - a) A box plot of activity time. b) Statistically significant correlations between the tasks (LaToza et al., 2006).

The results also revealed that most of the developers engaged simultaneously in more than one activity during the week with the majority engaging in all 9 activities all the time. The research also revealed considerable correlations, both positive and negative, between the various tasks, identifying the sequential nature of many of them (Figure 5.6 b). What was more important was the observation that the task distribution of developers changed together with the life-cycle of the project and the role of the engineer. The results in the figure clearly show a statistically significant correlation between the stages of the project and the main activity carried out by the developer.

The variation of tasks did not span across activities which are unrelated to software development (e.g. consulting activities). This suggests that the transition between the widely understood roles is not dependent on the life-cycle of the project. Indeed, in complex organisations, the distinction between Software Developers, Testers, Architects, Administrators, Consultants, etc. is much clearer and usually user progression through those roles takes time. As a consequence in the majority of cases user classification can be perceived as fairly stable.

This image has been removed

Figure 5.7: Task Distribution through the project life cycle - Proportion of different activities during different stages of the software project (LaToza et al., 2006)

The study carried out by Sharif et al. (2011) investigated the information-seeking behaviour of open-source programmers. The study was unique as it did not investigate a single developer-group but it compared the behavioural patterns of 6 different open-source groups. Moreover the results were captured over a period of several years, providing an insight into the developers' behavioural dynamics. The study focused primarily on analysis of the magnitude of information-need associated with different development tasks and problems. Figures 5.8 and 5.9 illustrate the results of the study.

The study identified differences between the development groups as well as changes in behaviour at different points of time. On the other hand the study identified very clear trends in software development task distribution. As illustrated in Figure 5.8, software developers engaged in IS&R primarily when working on implementation of software or when they were being introduced to a new technology. The results also illustrated that they did not seek information when working on the documentation, going through the change management process, testing, etc. (see Figure 5.8).

What is also very interesting, the need to use search utilities varied with the class of question developers tried to address. Figure 5.9 illustrates the distribution of the problems developers attempted to solve as well as the frequency with which those problems occurred. The results illustrated that developers engaged in IS&R primarily in order to confirm their understanding of the problem (i.e. "Confirmation" strategy) or when they were seeking for the solution of the problem-at-hand (i.e. "How" strategy).

The primary focus of the study, carried out by Ko et al. (2006), was to investigate

Figure 5.8: Frequency of software developer information-related problems - Sharif et al. (2011)

These graphs have been removed

Figure 5.9: Frequency of software developer information-related problems 2 - Sharif et al. (2011)

the mechanisms involved in acquisition of knowledge related to new code. The study involved 31 Java developers carrying out pre-designed development activity (re-design of a simplified paint program) over a period of 70 minutes. The user interaction was recorded using a screen capture software and annotated manually after the study. Additionally artificial interruptions were introduced during the 70 minute time interval. Figure 5.10 illustrates the results of the above process. The results showed that more than a quarter of entire study time was spend on textual search. Developers were capable of handling interruption very effectively. As illustrated in figure 5.11 they had very frequently swapped between searching, navigating and editing and the swapping process did not introduce significant overheads.

This graph has been removed

Figure 5.10: Developers' division of labour - Ko et al. (2006)

What is interesting in the context of this research is the observation that Software Developers, being the people directly involved in program delivery, used the Internet much more frequently than the more general group of Software Engineers. They used the Internet for significantly different purposes as well, including the use of the Internet not only as a knowledge source, but also as a source of snippets of various complexity and length, that are re-used in the process of opportunistic programming (Umarji et al., 2008). This makes this user group and the tasks of software delivery, tests and debugging, perfect candidates for further investigation. The selection of Software Developers for further research is not only based on the higher frequency of interaction

This image has been removed

Figure 5.11: Developers' work breakdown for two different tasks - Ko et al. (2006)

but also includes:

- The stability of Software Developers' Information Behaviour: Overall the literature is quite consistent in terms of the description of the Software Developers' work tasks. The research focusing around Software Developers' Information behaviour (Buckley et al., 2006; Freund et al., 2005) and the psychology of software development (Ellis and Haugan, 1997; Janjic et al., 2010) suggested that there exists a fixed and very limited set of work-related tasks / archetypes of behaviour that are being carried out systematically. In various user studies (Curtis et al., 1988; Freund et al., 2005; Gallardo-Valencia and Sim, 2009; Gallardo-Valencia et al., 2010; King and Griffiths, 1991; Seaman, 2002; Singer et al., 1997) identical tasks, micro behaviours and their distributions across various stages of the project were reported. User selection of activities and information needs was very rational and depended on the project temporary requirements.

- Role characteristics: The literature was also consistent in recognising the fact that various factors related to tasks and users do have impact on user information behaviour. The project role / project stage was one of the most frequently quoted characteristics of the task which was actually affecting the user performance. The results also suggest that the stage of the development project was significantly correlated with the set of software related responsibilities and as a consequence with the search requirements. The different roles (architectural, code maintenance, code testing and development tasks) required the usage of different data sources/search strategies. The temporal role of the developer also resulted in the different use of the captured information. In this context it is really important to select a single role, as the information requirements are significantly distinct across them. The code development activity is quite distinctive in terms of information requirements from other engineering tasks. What is also very important, is the fact that the reported software developers' behaviour across different stages of the project is consistent. Developers engage in the same tasks but change the proportion of time devoted to individual activities.

- User characteristics: Similarly the user expertise/maturity was shown to have a significant impact on the information choices and work-related behaviour. Users with significant domain expertise (consultants) focussed on much more abstract activities than novices. They often did not generate any code-related output. On the other hand, users with lower expertise were more often engaged in code delivery and tests and had much higher and quite re-occurring information needs.

- Limited number of development related sub-tasks: The literature very clearly illustrated that Software Developers were engaged in a very small and clearly defined sets of tasks. The main differences in findings were reported due to different expertise levels of the selected user groups. The experience of the participants of the studies mentioned in this review varied from the purely academic (Umarji et al., 2008) to the significant expertise of software consultants who were not engaged solely in software development (Umarji et al., 2008). If we focus on the common denominator of all of the groups, that is software related, software engineering tasks (Figure 5.5), the similarity of the findings and especially of the task distribution is very high and the quantity of tasks identified is very low.
- High commercial impact of Software Developers' activity: The existing literature acknowledges the need for the development of dedicated search utilities but unfortunately there exists a very small number of real implementations of domain specific personalized search or collaborative-filtering solutions (Janjic et al., 2010; LaToza et al., 2006). Even though a number of code-related search utilities exist, they were very rarely used due to their low efficiency, perceived reliability and usability (Sim et al., 1998).

In such an environment, where the information retrieval is executed so frequently and where the variation of tasks and contextual settings was very limited, software developers showed a significant potential for the development of very effective information-related habits / heuristics of behaviour. This, on the other hand, created an opportunity for the delivery of very effective, domain-specific, information retrieval/recommendation solutions (Ajzen, 2005).

5.5 Work analysis

In order to focus on the analysis of more specific IR related behaviours further work analysis has to be carried out. The work analysis has to be focused on the goals of the next methodological step: identification of IR behaviours which will be then used to optimise the user activity. The priority in the behaviour selection and the factor significantly impacting the analysis process is the amount of time spent on the particular activity. By focusing on the activities carried out most frequently it is possible to:

- Gain access to a significant sample of behaviours.
- Observe high levels of user "self-perception" (**?**).
- Observe habituation or behavioural automaticity (Fishbein and Ajzen, 2009).
- Achieve the highest levels of impact in terms of the observed change in interaction time and task efficiency.

The work analysis approach could be carried out using existing methodologies (Vicente, 1999) or a custom approach. Usually IS&R adopted a custom approach during the work analysis. In this research the work analysis was carried out based on the current understanding of the software developers' work and HIB. A significant amount of work was carried out on the psychological and cognitive aspects of software development (Détienne and Bott, 2002; Hutchins, 1996; von Mayrhauser, 1995; Xu and Rajlich, 2004). The research in Psychology of Programming (PPIG) (Buckley et al., 2006; Detienne and Bott, 2001; Hoc, 1990; Weinberg, 1998) and usability contributed significantly to the construction of development languages, colouring, naming and formatting guidelines, as well as the design of utilities supporting this cognitively complex activity. Modern IDEs significantly decreased the cognitive load required for the software development process by providing users with multiple tools automating or supporting the most repeatable tasks. Utilities such as on-demand and in-line documentation retrieval with support for obtaining code snippets or the availability of code analysis tools has contributed significantly to the improvement of the modern development process. Although significant improvements were made developers still spend a significant proportion of time on searching, retrieving and applying information to solve code specific problems (Buckley, Exeter and Good 2004).

Similarly in the Information Science domain there exists a vast body of research focusing on software developers (Case, 2007). When discussing the landscape of IS&R domain in 1984 Tom Wilson noticed that "the study of information-seeking behaviour

Web session intention:	Learning	Clarification	Reminder
Reason for using Web	Just-in-time learning of unfamiliar concepts	Connect high-level knowledge to implementation details	Substitute for memorization (*e.g.*, language syntax or function usage lookup)
Web session length	Tens of minutes	$\sim$ 1 minute	$<$ 1 minute
Starts with web search?	Almost always	Often	Sometimes
Search terms	Natural language related to high-level task	Mix of natural language and code, cross-language analogies	Mostly code (*e.g.*, function names, language keywords)
Example search	"ajax tutorial"	"javascript timer"	"mysql_fetch_array"
Num. result clicks	Usually several	Fewer	Usually zero or one
Num. query refinements	Usually several	Fewer	Usually zero
Types of webpages visited	Tutorials, how-to articles	API documentation, blog posts, articles	API documentation, result snippets on search page
Amount of code copied from Web	Dozens of lines (*e.g.*, from tutorial snippets)	Several lines	Varies
Immediately test copied code?	Yes	Not usually, often trust snippets	Varies

Table 5.3: Three points from the spectrum of Web Use intention

can be said to be the study of scientists" (what in the context of IR usually means engineers). The same is also valid for software developers. The way they interact with information on a day to day basis is currently widely understood (Ellis and Haugan, 1997; Freund et al., 2005; Gallardo-Valencia and Sim, 2009; Seaman, 2002; Singer et al., 1997). The rest of this section will focus on various work activity classifications and the distribution of developer activities.

Brandt et al. (2010*a*) carried out one of the most representative studies focusing on the impact of information objects on user IS&R. The discussed research suggested that the way software developers search for code can be classified on a linear scale depending on the novelty/intention of the information which was required (Table 5.3). The position of the information on that scale impacted not only on the way users searched for code, but also the core search characteristics/behaviours. Table 5.3 illustrates three example points on three extreme points of the spectrum and provides a brief description of the observed properties of the search process.

Similar classification was reported by Umarji et al. (2008), who also classified the web searching behaviour of coders' based on the same spectrum (code for re-use/reference example). This study used the size of the code-related content for classification of user behaviour. The study focused on the investigation of the overall distribution function of the code-related activities in this two dimensional space. The results of the investigation are presented in Table 5.4.

Sim et al. (1998) proposed a different approach which focused on the identification of the key activities in which the engineers were engaged. The activities/ habits were referred to by the author as "Archetypes" due to their wide re-occurrence across the

This table has been removed

Table 5.4: Purpose of search by target of the search. Numbers represent the number of observed search instances classified into the particular category (Umarji et al., 2008)

code searchers population. The questionnaire based user study was carried out on a group of 67 readers of the technology specific news groups. The identified tasks were classified based on the search motivation and were reported to span over the majority of software development activities. The most common search tasks reported were: defect repair, code re-use, program understanding, feature addition, and impact analysis. On top of this classification 11 archetypes of software development behaviour were designed and used for the in-depth analysis of the search process, search behaviour and the decision-making process.

The discussed research was then further extended to support the variation of developers' behaviour or use of archetypes across different stages of the project (Janjic et al., 2010). The results suggested that most of the specific information, related to the usage of code, library or framework was reported during the implementation and testing stages of the project. Other stages of the project involved the retrieval of more abstract information. It is important to notice that in this example the notion of Software Developer was used quite arbitrarily. In a significant proportion of projects, the phases of analysis and design, as well as deployment and maintenance would be assigned to different engineers in the company. The research shows (Figure 5.12) that the distribution of roles in the project even further narrows down the variation of tasks of people responsible for code delivery.

This supports the results from LaToza et al. (2006) which discussed the variation of the importance of different tasks across different stages of the development life-cycle. The notion of software developer archetype was also studied in the methodological and

This image has been removed

Figure 5.12: Task Distribution through the project life cycle 2 - Task distribution across different stages of the development process (Janjic et al., 2010).

organisational domain (Sawyer, 2004).

A very interesting study of information-seeking behaviour of women in IT was carried out by Choo and Marton (2003). The study involved 24 participants from 20 organisations. Study participants were asked to install dedicated software designed to observe their interaction with search utilities. After a period of automated observation, the participants were invited for an interview that was designed to extend the understanding of the recorded search episodes. One of the key outputs of the study was the analysis of developer search behaviour across different classes of "search episodes" (based on Ellis and Haugan (1997) classification). The results of this analysis are shown in Figure 5.13. The results reconfirmed the hypothesis that the task/ problem-at-hand has a non-trivial impact on the search strategy selected by the user.

The research carried out by Buckley et al. (2006) investigated the process of formulation of the search strategy during software maintenance tasks. This qualitative study involved two professional developers recording their work activities using a voice recorder for two hours. The analysis of the recordings allowed for formulation of the model of information-seeking presented in Figure **??**.

Software-developer behaviour seems to be consistent with the Principle of Least Effort (Zipf, 1949). It was shown that developers usually pay minimal effort to find the most optimal solution and intend to focus on the identification of the working solution meeting the requirements. In many cases, developers were not interested in

Figure 5.13: Distribution of information seeking episodes - Choo and Marton (2003)

These images have been removed

Figure 5.14: Model of Programmers' Information-Seeking Behaviour - Buckley et al. (2006)

the understanding of the problem or piece of code that was being implemented, but only keen to get a working solution (Detienne and Bott, 2001; Kim et al., 2004; Yeh et al., 2008). As a consequence software engineers often copied the solutions, even very loosely matching their problem and, "see what happens", engaging in the process of programming by example/iterative programming (Freund et al., 2005; LaToza et al., 2006; Philip et al., 2012). Lawrance et al. (2008) argued that programmers who were seeking for code, behave like information foragers. The study also showed that programmers spend about 35% of their time navigating source code. They behave like predators using basic instincts to follow the most promising leads of information in order to solve clearly defined problems and do not necessarily understand them in full detail. The ability to search, copy-and-paste and then experiment on the sections of third party code was very valuable to the programmers because it allowed them to distribute their cognitive resources across the Internet (Detienne and Bott, 2001; Détienne and Bott, 2002; Gallardo-Valencia and Sim, 2009; Hoc, 1990; Hollan et al., 2000; Hutchins, 1996; Weinberg, 1998). Owing to the availability of the on-line resources they frequently ignored the detail of the implementation of simple snippets of code what allowed them to focus on glueing together bigger sections and modules. The literature suggested that the extent of the code re-use through copy-and-paste was very high and reached up to 30% of the code in the solution (Gallardo-Valencia et al., 2010). The research carried out by Kim et.al identified that the average user performs approximately 16 copy and paste operation per hour (Umarji et al., 2008) out of which 2 are non-trivial sections of code and all of them carry relevant information. The research by Umarji et al. (2008) confirmed the findings above. The analysis of the development search activities showed that more than 73% of search sessions could be classified as opportunistic attempts. What is even more interesting is the fact that the proportion of non-trivial copy-and-pastes' were reported to be quite high and reaching up to 41%.

When the code examples were not available or the problem was more abstract and required an understanding of the underlying code, developers tended to look for other sources of highly processed information such as highly specialized discussion forums or blogs. Even when accessing such resources, the process of code delivery was in most of the cases similar. It was highly iterative and based on experimentation rather than an understanding of the problem. A very comprehensive piece of work on software code-

This table has been removed

Table 5.5: Classification of IR utilities based on their goal (Gallardo-Valencia et al., 2010)

cloning practices and their impact on development efficiency and quality was carried out by Kapser (2009).

The behavioural regularities reported above attracted a lot of interest in the Information Retrieval research community which led to the development of a significant number of research and commercially-focused search utilities (Gallardo-Valencia et al., 2010) as illustrated in Table 5.5.

The use of software-focused search and recommendation systems was very limited. The results suggested that in a majority of cases (87%) the general purpose search engine was used (Sim et al., 2011). Figure 5.15 clearly demonstrates that even though developers are aware of dedicated software IR tools such as Source Forge they do not use them regularly.

	Count
Google	28
Specific web sites	9
Mailing lists and forums	6
SourceForge	3
Scientific articles	2
Yahoo!	1
Krugle	1
freshmeat.net	1

	Count
Google, Yahoo!, MSN Search, etc.	60
Domain knowledge	37
Sourceforge.net, freshmeat.net	34
References from peers	30
Mailing lists	16
Code-specific search engines	11

Figure 5.15: Search System Preference - The exemplary distribution of preferred search sources. Left: open question response Right: Closed question response

There are multiple reasons why developers preferred to use a general purpose search utility:

- There was a high cognitive cost of engaging in a different way of searching and their limited scope, focusing almost exclusively on raw code. General purpose search engines allowed the user to search through various types of content starting from raw code to highly processed specialised articles. Moreover they allowed the user to change the preferences towards the information objects dynamically without the need to transfer to a different tool and repeat/ re-create the search process.
- The general purpose search engines allowed the user to be unspecific. They employed algorithms such as latent semantic indexing and probabilistic term selection to provide the user with relevant information even if the search phrases were not accurate or simply incorrect.
- Users had existing experience in the use of general purpose search engines and were expert in the formulation and reformulation of the search queries to suit their information needs.
- The majority of content stored by the specialised search utilities were indexed by the general purpose search engines.
- They employed much stronger relevance-ranking algorithms which provided the user with much higher accuracy results.
- The need for more complex sections of code, such as classes, methods and groups of methods, or descriptive information, inevitably led to the use of general purpose search engines. This is because they offered much higher levels of trust, the ability to control its behaviour and supported the user in expressing much more complex thoughts.

Overall the work analysis carried out in this part of the methodology suggested that, indeed, software developers interacted with information resources in a very consistent manner. They engaged in a finite number of information-demanding activities and required a finite number of information-object types on a day to day basis. Their key activities were very repetitive and could be frequently described through behavioural archetypes. Developers were not engaged in a single archetype of behaviour but in multitasking. The proportion of archetypes varied with the life-cycle of the project and with the project itself due to high cognitive load and time pressure. The extent of opportunistic programming also varied along with the problem and the information object required. As discussed above, there existed a wide range of software-specific IR tools and developers were aware of their existence. But unfortunately, they were not using those tools due to their various limitations.

5.6 Identification of behaviours: Behavioural consistency

The results of the work analysis as described in the previous section revealed a number of consistencies in the software-development related Information Behaviour. However this did not lead directly to the development of IR systems. In order to proceed towards the system development of the IR solution, the next step in the proposed methodology, "Identification of Behaviours", was carried out. This section illustrates how this methodological step was carried out in order to further narrow down the scope of the study. This was achieved through the minimisation of the number of variables/behaviours that were under investigation. This section focuses on the identification of the behaviours based on the following behavioural groups:

- User preferences towards information objects
- User preferences towards interfaces
- IR related behaviours

5.6.1 User preferences towards information objects

The previous section illustrated that even in the very narrow context of software development a variety of different information-problems occur. The problem variation led to varying information needs and changing requirements towards the information objects. The key information objects identified through the software developers' "work analysis" were as follows:

1. Software-related textual information: this included a variety of different information objects with varying reliability, complexity and level of abstraction. The user requirements for those objects were dependent on various factors:

   (a) The main factor affecting user perception of this information object was the size/scope of the underlying code structure. Depending on whether the user searched for blocks of code, subsystems, systems or frameworks the requirements for information will changed. This parameter directly correlated with one of the key attributes of the proposed methodology, that is, the perception of control. The smaller and more containable is the element, the more likely it is that the low quality information or information from a less reliable source (such as blogs, discussion forums and open source sites) will be used.

   (b) The more perceived control the user had over the action of selection and reuse of the identified information object the more likely the object was embedded in the development solution. This factor also directly correlated with the extent of social acceptance of the information reuse.

   (c) The perception of information also changed along with the reason for the IS&R. The information was used for multiple purposes including learning, clarification or in order to remind (i.e. distribute the cognition). Information objects did not have to be related to any specific development problem but could be retrieved because of curiosity or strictly for learning purposes. In such case their use was not directly related to any of the undertaken projects or tasks.

2. Code examples: these included a variety of different information objects also ranging in terms of size and complexity. The work analysis revealed that developers frequently looked for code for reuse. The retrieved code varied depending on a number of factors similar to the textual information objects:

   (a) It varied in terms of the size/scope of the underlying code structures. It could vary from the block level (spanning across single command and lines of code) through to the system level often spanning across entire systems (e.g. SharePoint) or frameworks (e.g. Linux).

(b) Finally, similarly to the textual information objects the code-related information varied in terms of the level of perceived behavioural control. As a consequence the level of behavioural control directly affected the frequency of observed code retrieval. It was highly expected that the retrieval of the block or sub system related code elements were much more frequent than the retrieval of system level code elements.

The use of code examples was a very important aspect of Software Developer Information Behaviour. The work analysis carried out in the previous section illustrated that more than 73% of all IR attempts could be classified in this category. Moreover the action of reuse of code could be successfully used for the identification of Implicit Relevance as the engagement in the opportunistic process, especially through code reorganisation and rebuilding, was an expression of code relevance. It was highly unlikely that the developer would paste into the software environment an irrelevant section of code and successfully rebuild it. Moreover, because of the very well-defined structure of code, it was fairly easy to identify it in the context of the web page. It was also very easy to classify it in terms of the technology used, its complexity, availability and size/scope. This task was even simpler in the context of the development environment.

As a consequence of the above analysis the block and subsystem level textual and code-related information objects were selected for further analysis with a significant preference towards code-related information.

5.6.2 IR related behaviours

The work analysis discussed the user interaction with the retrieved information and focused significantly on the tendency of users to reuse code-related objects in their software (Brandt et al., 2010*a*, 2009, 2008; Hartmann et al., 2008; Kapser, 2009; Philip et al., 2012). Such behaviour was very interesting from the perspective of IR system design because it had a high potential for being a very strong implicit-feedback indicator.

It is important to note that the reuse of code was not a distinct behaviour per se, but fell into one of many categories depending on the user attitude towards the behaviour, the rationale for its occurrence and the perception of control over the code reuse. As the work analysis illustrated developers engaged in code reuse to achieve the following goals:

- To distribute their cognitive resources and to use the Internet as an on-line repository of snippets that were not immediately available in the code at hand. The extent of this behaviour varied with the size of the information-related need (block, subsystem or the system) but mainly focused on the block level interactions.
- To engage in opportunistic programming which differed from the above with the user attitude towards the behaviour. Opportunistic practices did not relate to the reuse of standard and highly repeatable elements of code but to more abstract and complex code elements. The opportunistic practice mainly related to block and subsystem level code because the perception of control decreases as the complexity of the problem increases.
- To achieve component reuse: This sometimes refered to a deliberate (usually well considered) decision about the reuse of some other (often third party) open source or proprietary elements of the environment, but more often referred to more complex elements of the development process, such as subsystem or system level components.

The above mentioned behaviours were suitable candidates to be used in implicit-feedback implementations. They benefited from an ease of detection as the lexical classification of the search terms based on the key technological phrases was straightforward. The research studies suggested that they are very frequently observed and as a consequence would provide a significant body of behavioural data within a very short time. The behaviours were highly socially acceptable and the users, especially for the distribution of cognition and opportunistic behaviours, had a very high perception of control. The users perceived the behaviour as highly important to their daily activities.

In contrast to the above, the use of the purely textual information was not reported to be as clear. In reviewing the above studies a significant number of varying behaviours and activities relating to work involving textual search problem were identified. Across all search behaviours identified neither had a very high potential to correlate with relevance nor had a high potential to be used in implicit-feedback implementation.

The following behaviours were identified that have moderate correlation to information relevance:

- The direct use of the retrieved information when delivering code. The literature suggested that frequent attention-changes between the development environment and the IR system could be an indication of information relevance.
- The estimation of the probability of relevance could also potentially be increased through the observation of the user interaction with the copied elements of code, its compilation and debugging.

As a consequence the user code re-use behaviour and the above mentioned short-listed supporting behaviours were selected for further analysis.

5.6.3 User preferences towards interfaces

The work analysis in the above studies illustrated that, the users were not interested in the use of specialised search engines. Developers preferred the flexibility and efficiency of general purpose search engines which additionally allowed them to reduce the level of cognitive load that would be required by switching between the tools. The relevance algorithms used by the specialist IR solutions were limited and the overall cognitive cost of their use was very high, requiring from the user multiple changes to the search terms. The general purpose search engines offered developers the capability to use techniques such as LSI or probabilistic term suggestion in order to support their lack of knowledge.

Apart from the dedicated search utilities the work analysis revealed that there existed a limited number of search utilities which embedded themselves into the user development environment. They very rarely offered functionality significantly wider than the tools embedded into the IDE and because of the time delay they introduced were hardly competitive. Few systems offered the developer IS&R support for bigger information objects (sub-section, section). They were research focused and their validation and availability was limited.

The analysis of user preferences towards the interfaces carried out in this section was strongly driven by the selection of information objects and behaviours. As a consequence the interface analysis focused mainly on the user preference towards the code/software related retrieval utilities. Because of the reported frequency of interaction, which was very low for complex system-focused elements the further review focuses on block and subsystem level interfaces. The literature review showed a significant potential in the use of example-centric development search utilities (i.e. tools that help the developer to find code examples and effectively embed them in the environment). A number of existing research attempts were carried out which illustrated various UI concepts targeting the problem of development by example (Brandt et al., 2010*b*). The identification of user preference towards interfaces carried out in this section is focused strongly on the user attitudes towards their selection. Depending on the size of the problem the following preferences were identified:

- Block level problems: For the block level problems where the user intention was focused on the distribution of cognition or reuse of repeatable code the key preference was to embed the search utility in the development environment itself. This was because the users had a higher perception of control and were capable of expressing the problems more precisely and consistently when working with block level problems. The micro, block level, search problems, were already addressed with great success by tools such as "Microsoft InteliSense"(code auto-completion and documentation utility, embedded in Visual Studio). Those integrated utilities allowed users to do quick and very dynamic searches on the system commands and review the key documentation related to them. A number of solutions existed (Brandt et al., 2010*b*) which aimed to build on this paradigm to provide support for bigger block level information objects. The reviewed solutions were not focused on the code-related information objects but provide the user with the textual information only. The existing interfaces aimed to overcome this limitation, by providing embedded utilities that allow the user to identify the code-related objects in text and easily embed them in their solution. This was more complex than the standard copy-and-paste activity. As a result it did not provide incentive to use the alternative user interface. In contrast to the existing interfaces the proposed interface focused mainly on the block level search. The developers were capable of expressing their intent accurately. Additionally the copy-and-paste behaviour of other developers was known. Therefore it was not required to use the additional textual information to support developers. They were capable of assessing the relevance of the retrieved information based on the code only. The proposed interface addresses the limitations as discussed above and is shown in Figure 5.16.

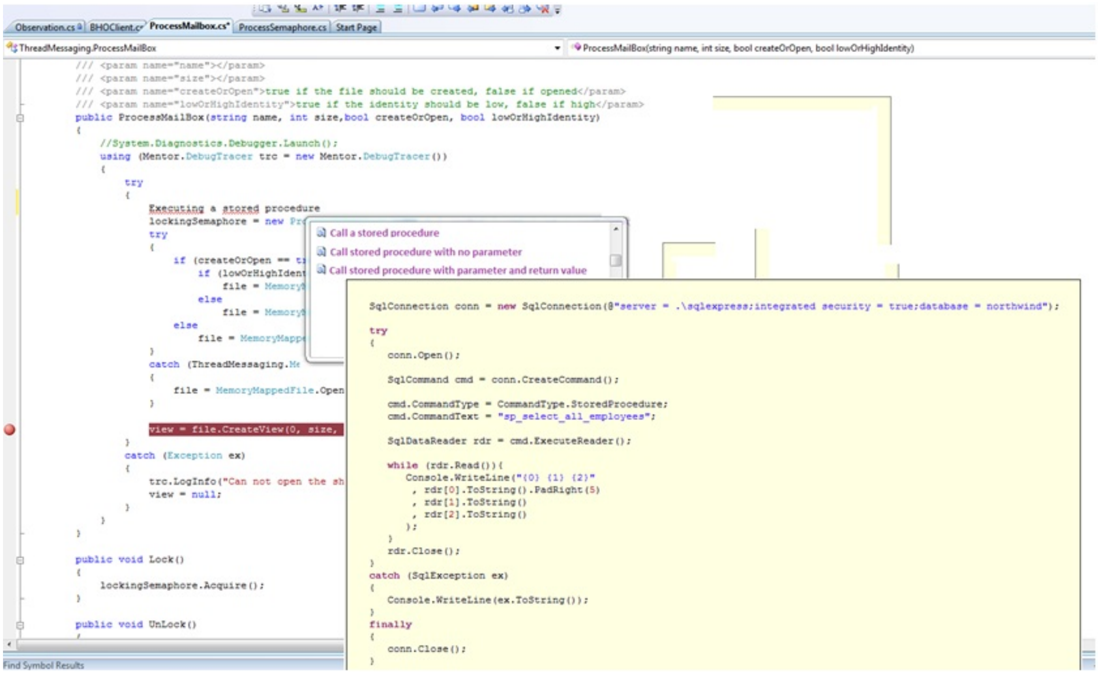


Figure 5.16: Preference Towards UI - The development environment focused user interface proposal

The proposed user interface needs to be efficient and ubiquitous so that it will not be distinguishable from the utilities already used. In other words the user should not know that he is searching the Internet when using the utility. This will remove the cognitive load from the user which would otherwise be required during the explicit search process.

- Subsystem and system level problems: The work analysis revealed a need to decrease the cognitive load by not providing the user with separate dedicated search utilities. Moreover it illustrated the catastrophic consequences of low retrieval accuracy which led to the users' abandoning the systems proposed. Therefore the natural consequence was to:
  - Provide the user with a UI taking advantage of the general purpose search

engine, for example by introducing a new, code focused facet of the information.

- Minimise the negative impact of low relevance estimation on the user perception of the quality of the system. This was achieved by minimising the cognitive load related to the assessment of the retrieved information as well as the number of actions that have to be taken to retrieve the code specific information. The proposed user interface which uses Google Search as an example (illustrated in Figure 5.17) was not designed to require any specific interaction from the user. It could embed into any standard Search Page (by the use of plug-ins). It also uses the same technique for cognition minimisation as the ones currently used by many advertisement providers. The interface embeds the additional content into the key page information in the regions of the secondary priority by the user (based on the understanding of user eye movement during search). As a consequence the code-related information is being-provided to the user all the time while the code-related problems are being detected and the user can simply ignore the information if it is of lower quality than expected.
- The proposed user interface has additional advantages as it offers the user the ability to control it in a consistent manner compared with the general purpose search engine.

The proposed user interface is focused on the retrieval of complex code-related information objects and expects the user not be able to recognise the relevant content immediately or solely based on a single facet of information. It offers the user multiple representations of the information as well as the ability to "scent" (Lawrance et al., 2008, 2010; Piorkowski et al., 2012*b*; Pirolli and Card, 1999) the information by clicking on the link provided.

This section discussed a systematic methodology to shortlist the information objects, IR-related behaviours and user interfaces (see Table 5.6).

This step was essential for the further narrowing down of the scope of work. The next section of this chapter focuses on the next step of the proposed 'Integrated Framework for HIB in-situ", that aims to identify the user attitudes and subjective norms

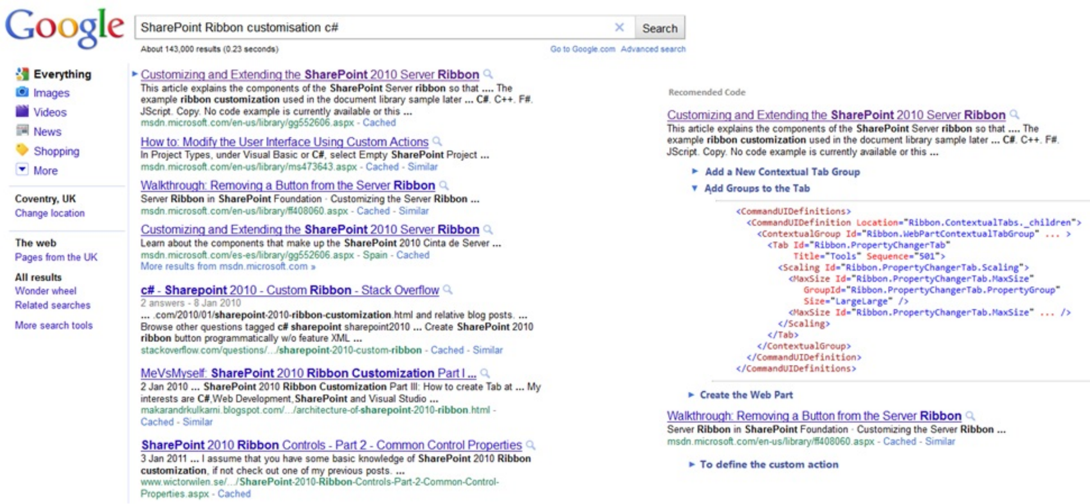


Figure 5.17: Preference Towards UI - General purpose search engine focused user interface proposal

Behaviour	Information Object Preference	Interface preference	Mechanism
Direct code reuse	Block Level	Development Environment	Distribution of cognition
Direct code reuse	Sub-system and System Level	Browser and Development Environment	Opportunistic programming
Direct code reuse	Components and Libraries	Browser	Component reuse
Indirect code reuse	Block, Sub-system and System Level	Browser	General software related retrieval

Table 5.6: Identified behaviours, preferences towards information objects and user interface.

driving the shortlisted behaviour. It also attempts to further understand the extent of behavioural control over the shortlisted behaviours.

5.7 Behavioural analysis

This section discusses the next step of the proposed methodology which is the "Behavioural Analysis".

The behavioural analysis was exampled by User Study 2. User Study 2 focused strongly on the three core elements of the proposed "Integrated Framework for HIB in-situ" including: attitudes towards the behaviour, subjective norms and the perception of behaviour control.

The aim of User Study 2 was threefold:

- To cross-validate the results obtained from the literature review.
- To ensure that the behaviours identified, in the previous methodological step will be observed with the sufficient frequency during the further analysis.
- To confirm that users perceive them as characteristic and re-occurring throughout the Information Interaction process.

Finally the goal of User Study 2 was to highlight the key factors that could moderate the observed behaviour so that they could be shortlisted for the analysis in the final step of the proposed methodology.

5.7.1 Procedure

The study consisted of two parts:

- A questionnaire that was emailed to the employees of the collaborating company. The questionnaire consisted of five open ended questions that asked for various aspects of user information needs, information behaviour and information use. The questions provided users with a number of selection options but also allowed for open answers which were highly encouraged.

- An interview that took place over multiple lunch times at the collaborating company by video conference. The goal of the interview was to gain additional understanding of the questionnaire results and gain knowledge about the geographical distribution of user work and the impact it has on their work patterns and Internet use. Additionally the interview focused on gathering real life examples of work-related Internet use, the associated information needs and the measures taken to satisfy them.

5.7.2 Instruments and materials

There were two main research instruments used in the study:

- A questionnaire that was distributed to the participants over the email.
- An interview question list that was used by the person conducting the interview to control the discussion and also take notes of user responses. The interview was recorded exclusively for the note taking purpose.

5.7.3 Data collection

The questionnaire consisted of two types of questions:

- Questions regarding the frequency of behaviours and interactions as well as their context:
  - Users provided qualitative information regarding data usage and perceived frequency
  - Users provided their comments on the context of information usage which were analysed. The context of information interaction included the information location, the task during which the interaction usually takes place and the conditions when it does not take place. In some cases the comments of the users affected the normalisation process of the qualitative information.
- Questions regarding the types of behaviour, information needs or information responses:
  - Close ended responses: allowed for collection of information regarding the shortlisted behaviours

ID	Age	Experience (years)	Gender	Key technology	Position
1	45	16	F	.NET	Support Developer
2	35	12	M	.NET	Developer
3	28	4	M	.NET	Developer
4	32	4	M	.NET	Support Developer
5	29	2	M	Java	Junior Developer
6	27	5	M	.NET, Java	Developer
7	27	4	M	.NET, Java	Developer
8	25	5	F	Java	Developer
9	25	6	M	Java	Developer
10	27	4	M	.NET, Java	Support Developer
11	43	15	M	.NET	Developer
12	31	12	M	.NET	Developer
13	32	13	M	.NET, Java	Developer
14	34	12	M	.NET	Developer
15	27	5	M	.NET, Java	Developer
16	26	3	M	Java	Junior Developer

Table 5.7: User Study 2 population demographics

- Open ended responses and comments: allowed the understanding of further the conditions of IR and the identification of other frequent behaviours not identified in the initial review. The results of the interview were analysed qualitatively on a question to question basis.

5.7.4 Experimental population

Table 5.7 summarises the key characteristics of the population of User Study 2.

5.8 Identification of moderating-factors

The investigation of the moderating-factors and their impact on the observed decision-making is the key element of this research. This study, User Study 3, supports the final step of the proposed methodology. The goal of User Study 3 is to identify how the shortlisted behaviours change along with the moderating-factors. The study focuses on the key factors identified in User Study 2 including:

- User context: understood as a re-occurring user specific combination of geographical location and time

- User task: understood as re-occurring observations of the domain interaction
- Type and size of information objects required
- The overall complexity of the problem
- The project with which the user is engaging and its age

User Study 3 also revisits the discussion related to human overall information consistency addressed in User Study 1. The analysis of HIB consistency is carried out in the context of software development. User Study 3 attempts to revalidate the initial observation regarding the distribution of human behaviour across the listed moderating-factors. It also focuses significantly on the overall consistency of user micro, session level and macro behaviour, further supporting the initial assumptions of human consistency and rationality.

In contrast to previous two studies, User Study 3 is significantly more quantitative. It consists of a long term automated analysis of developers' interaction with information resources. User interaction with the browser, the development environment and the operating system are captured by the experimental software and the results are based on the numerical analysis of this data.

User Study 3 concludes the methodological process outlined in the previous chapter by:

- Supporting the thesis of Developer Information Behaviour consistency
- Supporting the selection of the identified behaviours for the development of IR systems and identifying moderating-factors affecting their efficiency

5.8.1 Procedure

The experiment was composed of two stages:

- Questionnaire: The goal of the questionnaire was two-fold:
  - First of all, to determine the participants' suitability for the experiment. This step is essential due to the limitations of the experimental software which was capable of observing only a limited number of browsers and development environments. The selection of participants had to be based on

the users' choice of development tools, browser and pattern of work, as well as basic aspects of the user profile.

  - The secondary goal of a questionnaire was to determine the basic characteristics of the user population (such as technology experience, etc).

- Automatic observation of behaviour: During the second stage, the installed software captured user behaviour without the need for user intervention (unless a new environment is created or installation is needed or an update to the software is necessary). The data captured by the software was being stored on participants' machines and when Internet connection was available it was transported over the secure channel to the university server.

The process of installation, questionnaire distribution and collection was spread throughout a significant time span and each participant was also approached individually to minimise the impact on the projects carried out by the company. Therefore once a particular employee expressed his interest in the participation in the experiment he was approached individually, usually during lunch time and asked to fill out the questionnaire. Once the questionnaire was filled out, it was immediately assessed in terms of the participant's suitability, and if the participant was suitable the software was installed on all of his development environments.

One of the key concerns of the study was the extensive use of virtual and remote environments and their often rapid rotation. It is not unusual for the developers to have more than one development environment, each dedicated to a particular project, technology or customer. Each of those environments is usually absolutely isolated from one another using the virtualisation technology, each one having its own operating system and development tools. It was essential therefore to ensure that the user migration to a different development environment is both visible and is managed accordingly. To make sure that the process of migration is not obstructive to the user from the experimental perspective, every participant was provided with his own version of the installation file that already had the user credentials preconfigured. The package was delivered as a "One Click" installation and it was sufficient to provide the users with the URL of the installation file that had to be clicked in order to deploy the experimental software.

To ensure that the migration of the user is not unnoticed, the "SQL Reporting Services" were deployed on the experimental server and appropriate reports of user activity were available on demand. This allowed for the identification of time intervals when the development process was not being observed by the software which led to communication with the particular participant who was then prompted to have the software installed.

Once the experiment was over the software was removed manually using the control panel.

5.8.2 Instruments and materials

The user study 3 took advantage of two types of materials:

- Questionnaire and associated documentation: The questionnaire consisted of a number of close ended tasks aiming to identify user eligibility for participation in the study.

- Automated monitoring software: The software was embedded into the key programs that the user used in the workplace and carried out an automated observation of his interaction. The software did not have any user interfaces (apart from the credential and tracing screens not normally accessible by the users) and did not require any user interaction.

5.8.3 Data collection

The experimental software was composed of multiple components that focused on various forms of user information interaction. The observed interactions of the user can be classified based on the monitored program:

The raw information listed above was initially stored on the local micro database that was installed along with the experimental software. When there was sufficient observations gathered and when the Internet connection was available and not significantly loaded, the bulk reports were send to the appropriate user server over a secure channel. After the data regarding user activity was saved on the server a dedicated server process would trigger and monitor the URLs of the reported pages. The content of the Internet pages was then downloaded. It is important to notice that the experimental software did not capture the content of the pages on the user machines, the keyboard entry of the users, exact mouse movement, clicks, and session or cookie

Data captured in the browser (IE, Chrome and Firefox)	Data captured in Visual Studio	Data captured by the general program OS wise
- Timestamp when the page is opened - Timestamp when the page is closed - Active reading time of the page - Amount of scrolling - Amount of clicks on the page - Browser type - Domain name of the page for classification of content as work / non work related - GUID of the interaction - GUID of the previous interaction (allowing to analyse the entire session and not only the individual pages) - Amount of focus changes - URL of the page - As a consequence of the URL the search activities and queries - As a consequence of the URL the page content (under the condition the page is public) - IP of the requesting host	- Solution identifier - Project identifier - File identifier - File membership in a project and solution - Timestamp when the file is opened - Timestamp when the file is closed - Amount of scrolling - Amount of clicks - Number of build actions - Usage of the refactoring tools	- Copy and paste pattern determining from which sources the data was copied and finally where it was pasted (focusing mainly on the browser) - Focus change pattern determining how the user changed focus between the different types of information (focusing mainly on the browser and development environment) - IP of the machine hosting the program

Table 5.8: The data captured by the experimental software in User Study 3

information. As a consequence the input of the forms filled by the users remained anonymous and the system did not have access to the sensitive information. Similarly, since the crawling of the web was done outside of the network, the user, company or customer internal information was not accessible. The crawler indexed the public content only.

Once on the server the data was being extracted from the database using the dedicated SQL queries and exported to several different formats depending on the goal of the further analysis.

- A series of files was created to enable the consumption of the captured data by the TMG toolbox for Matlab. This step also included implementation of custom Matlab routines that allowed for consumption of behavioural data.

- The data was also exported as a number of processed CSV tables that were further consumed by excel and PASW for basic statistical reporting.

- A number of data transformation scripts were created to allow for session level analysis and consumption of information in a format required by an Excel extension for Graph analysis carried out with NodeXL.

Overall a significant number of data aggregates were calculated based on the data including:

- Interaction session information: based on the GUIDs assigned to particular pages could be captured. This information focussed mainly on session characteristics such as session shape, size and classification and session average reading time.
- Development tool interaction
- OS interaction
- User characteristics: extracted from the questionnaire
- Finally a number of data aggregates were created based on:
  - Domain of the content viewed
  - Work / private activities (manual annotation)
  - Classification of software development work-related activities
  - Time
  - Visual studio solution and project

5.8.4 Experimental population

Overall 21 users have responded to the advertising email and 12 were found eligible to participate in the study. Since the selection of the participants was carried out based on the same developer population as in User Study 2 the population has the same skews and identical characteristics.

5.9 Conclusions

This chapter discusses how the proposed methodology was applied in this research. It discusses all of the methodological steps and how they were applied in this study. The chapter describes how the key decisions of the study were made, regarding the selection of the population, work and search task as well as the selection of the key behaviours used for further analysis. The chapter also introduces User Studies 1, 2 and 3. It discusses the role they play in this study, the part they play in the proposed methodology and how they support the behavioural analysis process.

Overall this chapter demonstrates how the proposed "Integrated Framework for HIB in-situ" can be applied to the real life problem. It discusses the process end to end, illustrates how the steps of the process can be designed and what should be the expected outcome. In the context of this research the chapter demonstrates:

- The methodology for experimental assessment of general HIB consistency
- How the initial population, work and search task was selected: in this case short-listing the development related problems of software engineers. Moreover, it demonstrates how this selection can be supported by evidence, justified and initially assessed in terms of further work feasibility.
- How the work analysis of the particular user group targeted at a particular task was carried out. The work analysis focuses on the key types of behaviours: preferences towards information objects, preferences towards interfaces and other behaviours supporting design of IR systems.
- How the final behavioural selection can be systematically carried and what level of justification can be expected.
- How the behavioural analysis can be performed investigating the perception of control, attitudes and subjective norms.
- How the variability of the identified behaviours can be systematically measured using the results of the analysis. Additionally, how the knowledge of variability can be used to control the behaviour of the proposed IR systems.

The following chapter will discuss the results of User Studies 1, 2 and 3 supporting the discussion carried out in this chapter with empirical evidence.

Chapter 5: Summary

Human behavioural consistency is a necessary condition for the application of the ”Integrated Framework for HIB in-situ”.
The existence and nature of HIB consistency is verified through User Study 1. User Study 1 investigates the consistency of the parameters of the General Black Box Model (Perception of environment, Perception of task, User State and Human Information Behaviour) and the extent to which information-related habits are developed during IS&R.
User Study 1 involves 12 participants and consists of an opening and closing interview as well as a diary study.
The proposed methodology (”Integrated Framework for HIB in-situ”) is applied to the information problem of software engineer information behaviour.
The way particular activities of the proposed methodology are carried out is described. This includes Actor/Work/Task analysis, identification of the key behaviours, behavioural analysis and identification of moderating-factors affecting the selected behaviours.
The review shortlisted work-related software developer activity for further analysis. Further work analysis led to the identification of the key behaviours that are considered for further analysis (see Table 5.6).
In the context of the shortlisted behaviours the design of User Studies 2 and 3 is explained. User Study 2 directly addresses the Behavioural Analysis step of the proposed methodology. User Study 3 addresses the final step of the proposed methodology: Identification of moderating-factors.
The results of User Studies 1, 2 and 3 along with the result interpretation are presented in the next chapter.

Table 5.9: Summary of chapter 5

6

Results

6.1 Introduction

The previous chapter discussed how the proposed "Integrated Framework for HIB in-situ" was applied to the selected population of Software Developers. The previous chapter opened with the description of User Study 1, designed to investigate the problem of overall HIB consistency. It then continued through the steps of the proposed framework, supporting the key decisions of the research with literature review and introduced User Studies 2 and 3. It discussed the design of User Study 2 supporting the behavioural analysis step of the framework and concluded with the description of the design of User Study 3 supporting the final step of moderating-factor analysis.

This chapter discusses the result of above mentioned user studies. The further sections are focused directly on User Studies 1 to 3 and correspond directly to the key research questions that are:

- When and to what extent HIB is consistent ?(User Study 1)
- How the design of user interfaces, implicit-feedback algorithms and personalised IR solutions can take advantage of the consistent episodes of information interaction.
    - Is it possible to effectively identify the consistent behaviours? (User Study 2)
    - Is it possible to effectively identify the factors affecting the consistency and control the level of predictive validity? (User Study 3)

6.2 Overall HIB Consistency

User Study 1 was designed to investigate the extent to which humans are consistent when interacting with information resources. This section focuses on the results of User Study 1 which illustrate the extent of HIB consistency and the factors affecting it. The analysis of results is divided into two sections:

- The discussion of the overall HIB consistency in the light of the Black Box Model. The analysis is focused on the Black Box Model input and output parameters and aims to prove that:
  - Humans tend to choose reoccurring/consistent locations /contexts /environments when interacting with information.
  - Human interaction with information is strongly task driven and the task distribution is very consistent in time, space and across similar user populations.
  - The consistent distribution of tasks drives users to interact with information across reoccurring/consistent moments in time.
  - When working on reoccurring tasks in consistent locations and at the same points of time, user HIB is more consistent than the total HIB.
- The investigation of the tendency of users to develop habits/automate their behaviour when interacting with information.

6.2.1 HIB Consistency

6.2.1.1 Location Consistency

The perception of the environment is one of the key components of the discussed Black Box model of HIB. It shapes the way we perceive the world and it drives the way we interpret our activities and prioritise the required information. The location of the users when interacting with information was captured during the first two stages of the User Study 1:

- The opening interview: The users were provided a map and asked to pinpoint the key locations where they use IR systems. The map was then further attached to the diary itself. The goal of this exercise was mainly to get a general overview of the geographical distribution of user work and to gain a commitment from the users to describe all of the places where interaction takes place rather than where it was the most comfortable to get the diaries filled.

- The diary study: Each entry in the diary required from the user a selection of one of the locations previously identified during the interviews, or a description of the new location.

The results from the diaries and the interviews were merged together to gain an understanding of time and location distribution for particular tasks. Table 6.1 provides an overview of the location distribution captured in User Study 1. A big proportion of the locations reported in the interviews related to sporadic interactions which occurred very infrequently and as a consequence were not representative. Only the locations where the average observed, or self-reported interaction frequency was higher than seven times a week, were further analysed.

At early stages of the experiment it was noticed that because of the size and weight of the computer hardware users tend to be quite mobile within the locations of interaction. Experiment users tended to use the IR solutions in more than one place within their homes and in more than one place in their work places. It was clear that the selection of location was purposeful and that it should be introduced in the further analysis.

The values in Table 6.1 demonstrate that the majority of the user's everyday activities were carried out in the same locations. The variation of location is marginal and occurs mainly for low frequency tasks and information activities. When assessing the activities in the diary the users were also asked to quantify the frequency with which they carry out the reported activity, if it was reoccurring. The reported activities were then grouped by the places of interaction. Every section of a bar in Figure 6.1 illustrates the frequency (times/week) of a certain activity carried out in a single place of interaction.

The data presented in Figure 6.1 reveals a number of properties of user IS&R location distribution:

- It illustrates that only a marginal number of tasks are carried out with high frequency and that they are carried out very systematically at the same locations.

Participant	Diary Data		Merged and trimmed data where diary captured more than 7 interactions/week	
	Number of locations of interaction	Number of places of interaction	Number of locations of interaction	Number of places of interaction
1	2	4	1	2
2	3	4	1	2
3	3	4	2	2
4	4	5	3	4
5	4	5	2	3
6	2	3	2	2
7	2	4	1	1
8	2	3	2	3
9	4	5	3	3
10	3	3	2	2
11	3	3	1	1
12	5	5	1	2
Average	3.08	4.00	1.75	2.25
SD	1.00	0.85	0.75	0.87
Total	37	48	21	27

Table 6.1: The number of geographical locations, and places within those locations where participants interacted with information

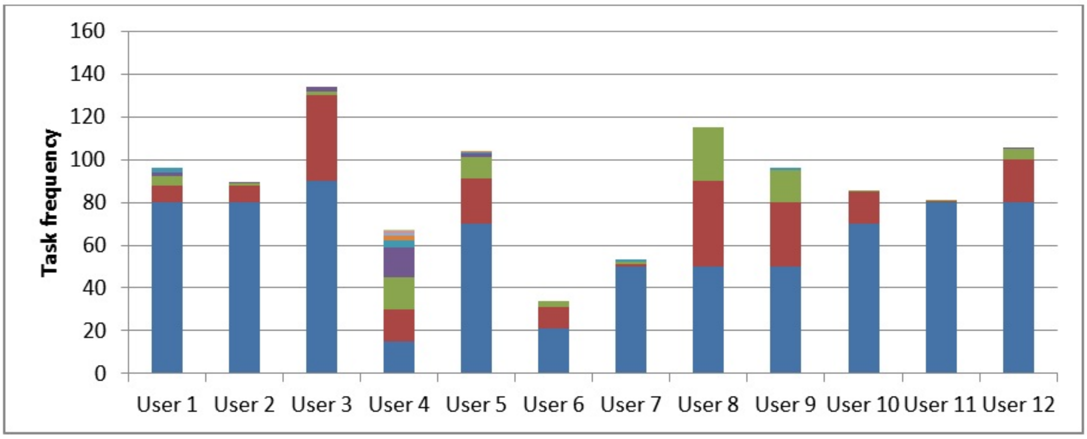


Figure 6.1: Interaction location frequencies - User task location frequency distribution

- It illustrates that the distribution of location selection frequency is highly non-uniform. The frequency of activities is not the same but a very small number of activities and locations dominate, with the frequency distribution resembling exponential distribution.

- It illustrates individual differences in the overall frequency of interaction with IR tools and the distribution of frequency of interaction.

The above observations are also very consistent with the information provided explicitly by the participants throughout the interview part of the study. Neither of the participants had any doubts when pointing out the places where IR occurs and their self-reported location distribution was consistent with the information reported in the diary part of the study. Participants also found it easy to quantify the average amount of time they spend in various locations using the IR utilities. The participants were also very effective in classifying the locations as routine or occasional and in a majority of cases were very proactive in making this distinction.

In contrast to the above, the rationale for selecting a location and engaging in the task at a certain place is reported to be much more complex. The decision about location selection is made based on a number of task, context and personal factors as identified in the following examples:

- Participant 1 reported leaving the laptop at home if on an engaging lecture or if there are no substantial breaks during the day. The laptop would be taken to the university and the Internet accessed for:

  - "Boring classes" - but only for activities not requiring a lot of attention such as Facebook,
  - Group work - and in that case the work is highly task driven,
  - Breaks in between classes - which are longer than an hour.

- Participant 3 reported an attachment to the usual place of work with only two very carefully selected locations where search activities could take place. The presence of the private space, private set-up and accessories is the primary factor affecting the selection of work location.

- Participant 4 was highly flexible and adjusted to the requirements of the task. This participant was able to carry out primary activity regardless of the circumstances although preferred certain places over other others for some of the activities.

- Participant 5 primarily valued the presence of other people and as a consequence would select locations where companionship could be found. This participant isolated himself only when engaged in last minute preparations for exams.

Similarly a variation between the participants' location choices was reported. Participant 3 reported that the "majority of search was done at home or at work". In the case of this participant the laptops would not be even moved from the usual place of interaction. All other interactions were reported to be sporadic, usually music and news focused, and carried out when visiting friends and family. To contrast that, some of the participants reported high flexibility in terms of the distribution of Internet use, reporting frequent movement and changes to work style across diferent locations (users would use a veriety of devices, interupt the primary activity do diferent extent, use diferent comunication resources, etc.). Participant 4, for example, reported having a very structured approach to the use of the Internet across multiple locations. She identified 5 main tasks that were being carried out at diferent locations, including work; university; university library; private room in the shared house and the common room. Participant 4 was very consistent in terms of task distribution and individual activities were always carried out in the same locations. In contrast to this Participants 6,8 and 9 reported a number of locations where they carry out all forms of search activity (private and work-related).

Overall, the participants clearly demonstrated, that the decision about the search location is not accidental. It is primarily task driven, but factors such as: perception of comfort; personal ability to concentrate; availability of miscellaneous resources (such as a TV set, printer, books, speakers, sofas); or simply other people, had frequently a decisive impact on the location selection.

The results also illustrated differences between the participants not only in the type of location,number of locations and frequency of use, but what is more important, the reasons why they chose them.

6.2.1.2 Time Consistency

The User Study 1 focused on the consistency and time distribution of user Internet activity. The primary interest at this stage was to understand the extent to which people carry out the same tasks at the same time consistently, how this distribution varies with time and task and what are the diferences in time distributions between participants of the study.

User Study 1 was composed of an interview stage and a diary study, both recoding the time related information. Since the data was captured in a diferent way during those two stages the analysis of this information is done seperately.

During the interview participants were presented with a weekly view of the calendar. The calendar was used durign the conversation to create a detailed record of user weekly interaction. It helped the users to focus on their habits, repeatable tasks and their motives for the reoccurring engagement.

The interviews revealed that users have a very strict time routine which was driven by their task and location distribution. This user focus on the task at hand and functional aspects of interaction was reflected in the way participants responded to the question about the time when they usually interacted with the Internet. None of the participants, when asked the question about the time of interaction replied directly by providing the time estimates. Instead, all of the participants had a tendency to disassemble their work pattern into functional chunks and describe them based on the properties of assignments and the requirements of the contexts of interaction. To illustrate the above mentioned phenomena, the data analysis of User 1 behaviour is shown below. It is then followed by additional aggregate information analysis to underline the common trends observed across all of the participants as well as the key differences between them.

In answering the question about interaction with IR tools, User 1 divided the working week into the following five classes of interaction which diretly mapped to the following location selection:

- T1 Internet use during the university classes: "But only if the class is boring and if it is not the last class of the day after which I go home".

- T2 - Internet use in the university studio rooms: "Every two weeks when there is group" and T3 - "during the longer breaks between classes as it makes no sense to go home because of the distance".

- T4 - Internet use at home in the bedroom: "Exclusively when there is university work to be done"

- T5 - Internet use at home in the common room: "During meals or when doing something for pleasure"

Figure 6.2 shows a histogram of User 1 daily-activities throughout the week as captured during the interview. Every bin in the histogram represents a single hour within the week starting from Sunday.

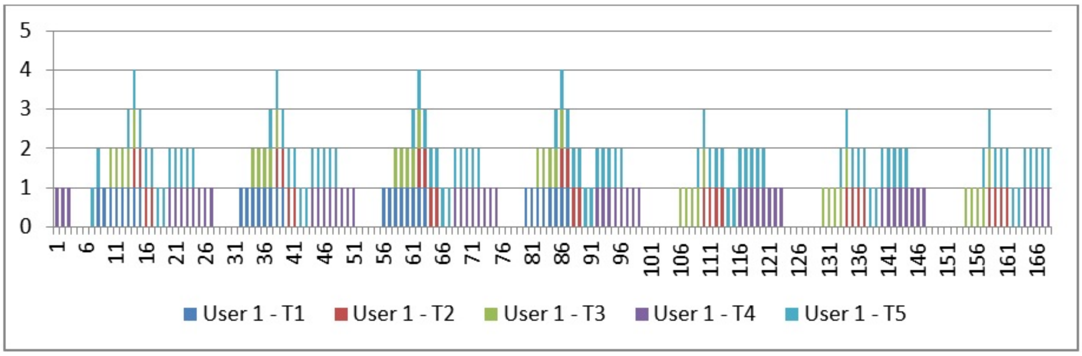


Figure 6.2: Calendar based User 1 activity time distribution -

This is a reoccurring trend in that none of the participants was capable of answering the question of IR interaction without analysing their task and location distribution. The user time distribution is task and context dependent. It means that the overall time distribution of Internet activities should be a combination of the independent random variables associated to those tasks, forming a mixed probability model of time distribution.

In order to verify the above, a more detailed analysis of user-activity distribution had to be performed. The data illustrated in Figure 6.2 could not be used for distribution analysis because the probability of a task occurring in the hourly time slot was not equal for all of the tasks but varied considerably. Because of that reason, the users were also asked to provide an estimate of task frequency, if it was reported to be reoccurring. Since every task reported in the calendar was had a different frequency, every occurrence of that task was multiplied by the self-reported frequency measure, producing the time distribution presented in Figure 6.3.

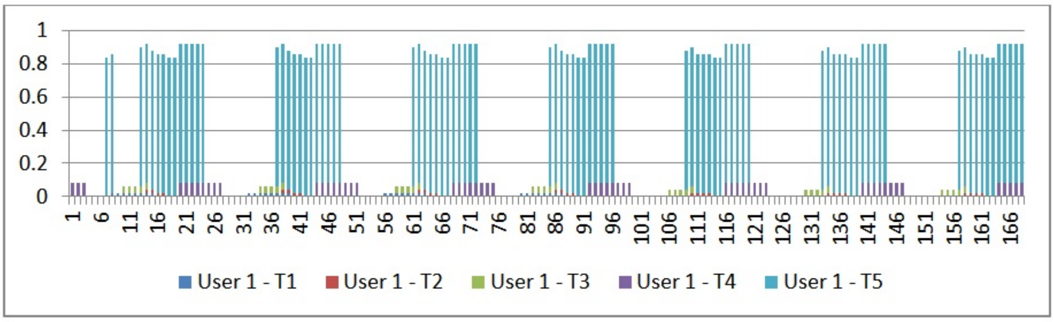


Figure 6.3: Calendar based User 1 activity time based distribution normalised based on task frequency -

Figure 6.3 clearly illustrates that there exists a certain finite number of tasks that will have a dominating effect on the overall mixed-model and clearly shows the marginal impact of side tasks on the distribution of user time interaction. This relationship is a characteristic of all the users of the study(Figure 6.1) even though the variations between the users in terms of the locations and the time distribution itself are substantial.

The observations above were further supported by the data collected in the diary study. Every diary entry was processed and the information about its start time and duration was extracted and merged with the interview based information. The information regarding the re-occurrence of the diary entries was not used. An attempt was made to identify the deviation of diary data from the interview data. The results of the analysis show an overlap between the data captured in the diaries and the self-reported distribution captured during the interview. Not a single deviation from the self-reported pattern was observed in the diaries. This is not a surprise as the users deliberately introduced an error margin into the information reported verbally during the interview. Regardless of the above this illustrates high levels of users' self-awareness and consistency between the behavioural attitudes (self-reported time distribution) and the actual behaviour (data captured in the diary study).

The overall shape of the time distribution obtained from the interviews was compared with the diary study. Figure 6.4 illustrates the analysis of IS&R time distribution of all participants gathered through the diary study and the interviews. The analysis of the histograms showed individual differences not only in the magnitudes of the histograms and their location but also in terms of the mixed-model itself. The histograms also supported the hypothesised mixed-model nature of user time distribution.

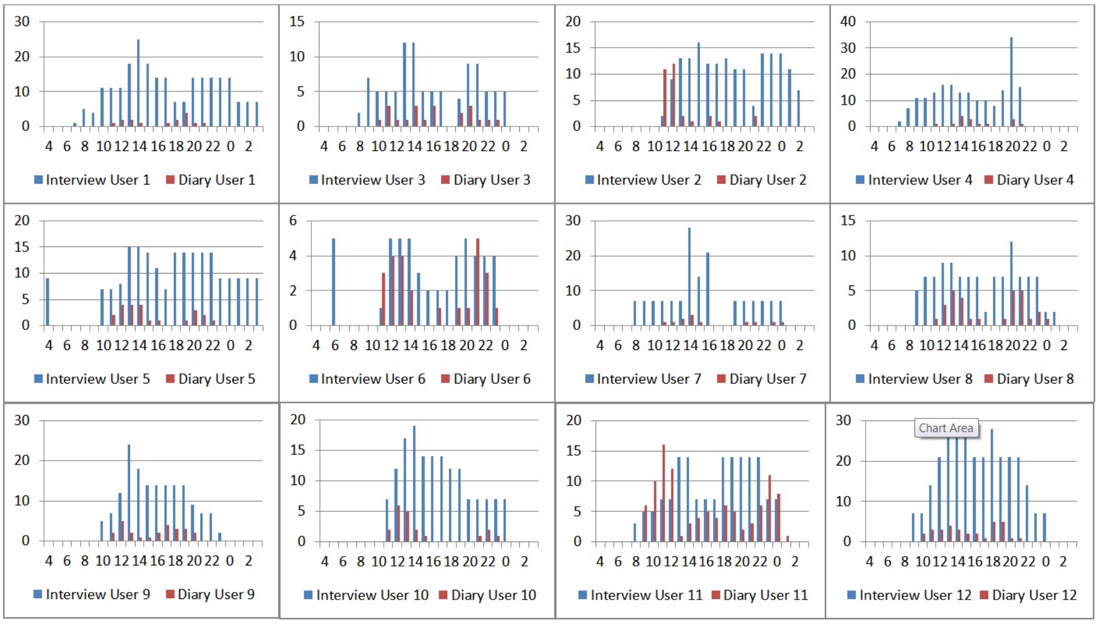


Figure 6.4: User Study 1 time distribution of all users - Histogram of User Study 1 participants' IS&R time distribution

This pattern reoccurs across the entire participant population and this is shown in Figure 6.5. Figure 6.5 illustrates the normalised data-aggregate histogram which highlights the impact of work and non-work tasks on the user interaction distribution. The histogram clearly illustrates a division of the time distribution into two phases of interaction:

- The first: occurring during the first parts of the day (not necessarily in the morning) and
- The second: being in the later stages of the day or night.

Further analysis of the histogram (Figure 6.6) showed that a mixed-model based on two normal distributions ($\omega_0 = .62, \mu_0 = 10.85, \sigma_0 = 7.24, \omega_1 = .38, \mu_1 = 18.32, \sigma_1 =$

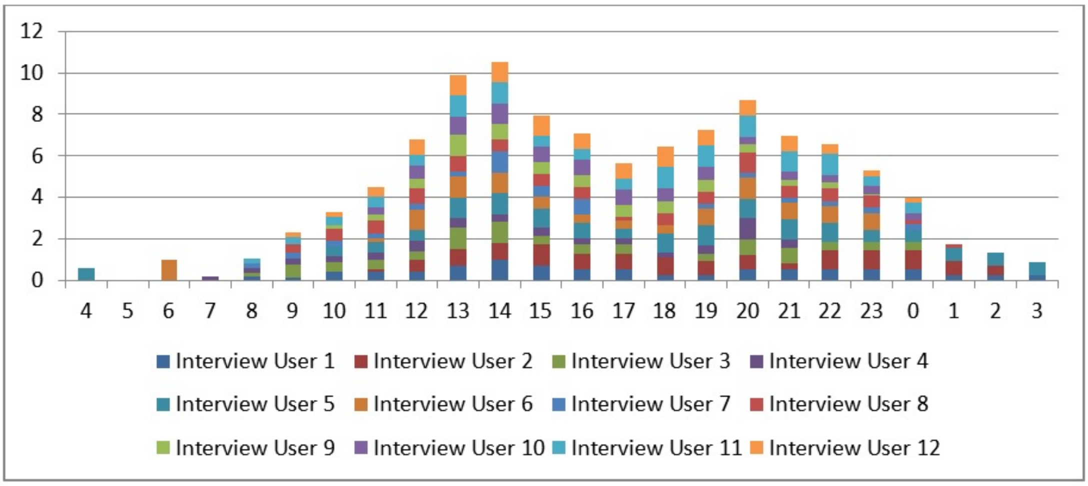


Figure 6.5: Normalised user-activity distribution -

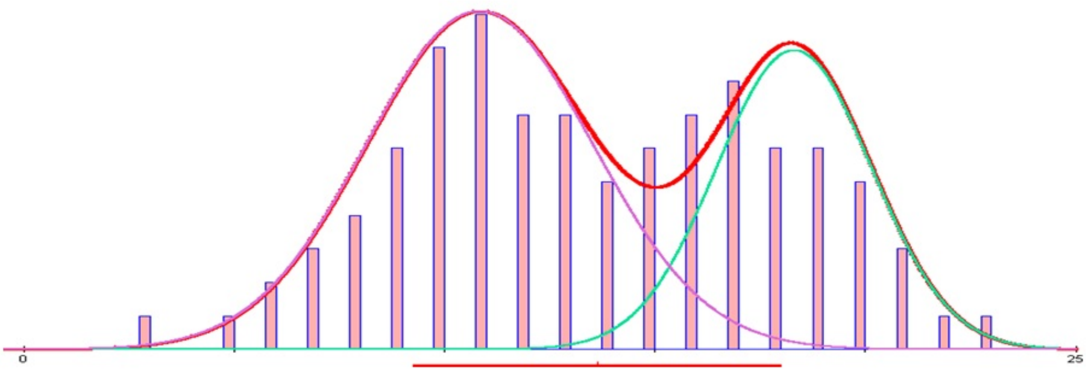


Figure 6.6: Bimodal nature of user-activity time distribution - Figure presents a histogram of the cumulative activity time distribution along with the approximation of two normal curves of the bimodal distribution.

3.41) could be created and that differences between the two normal components were statistically significant (z=2.32,p=.01).

Of particular importance was the fact that the diary study provided the same time / location task distribution as the interview where self-reported data was collected. This indicated that there was not only consistency in user activities but also that users were aware of it. Moreover, this suggested that such stabilisation was deliberate and that the users actively attempted to achieve it.

The above was supported by the verbal response gathered through the interview session which asked the question about the rationale for choosing the particular location. In all of the cases participants were able to clearly and easily support their selection with an argument, for example:

- User 1
  - Use of the Internet during classes: 'The computer keeps me awake when tired, bored or if it is early in the morning.'
  - Group work in the university during the afternoons: "When there is a group work then it is used for group work, once per session it is email facebook or cars."
  - Work in the bedroom: working there because of the space, comfort and privacy. "Highly dependent on the coursework and changes with the task. Really depends on what has to be done. A lot of research on the topic and then implementation. Using primarily sketching and drawing sites as well as TopSpeed.COM or flynews"
  - Work in the common room in the house: 'Because it is the most comfortable available place'
- User 4
  - Work in the library during the day: "Distance to home is big and I go there between the classes. In this place it is easy to focus and I try to do as much work as possible then as I need to go to work after classes."
  - Working in the library in the evening and over weekends: "It is easy to focus and there are books and printers there."

– Working in the library for group meetings: "Close to the main building and space and resources available. Other places are chosen only when all of the rooms are booked."
– Working in the kitchen: "This is a comfortable place but only for assignments. Sometimes I need to leave as people want to cook there."
– Working in the bedroom: "I work there only if I cannot work in the kitchen or the living room."

This shows that even though the decision-making process was usually quite complex and based on highly subjective and often emotional factors the responses of the user and their time distribution was very logical and as a consequence consistent.

6.2.1.3 Consistency of other moderating-factors

During the diary study participants were asked to subjectively assess 8 factors potentially impacting their IS&R behaviour (listed in Table 6.2). These factors focus on the key psychological moderating-variables associated with the current task. The goal of this analysis was to identify to what extent the moderating-variables changed across different tasks and geographical locations and to compare the extent of this variation with the variation on a per-user basis. The calculation of the Coefficient of Variance at per user and per population level requires the division of the standard deviation of the population, by the mean. The calculation of contextualized coefficient of variance was based on the oversimplified definition of context, which for the purpose of this analysis was defined as:

> "a user-specific combination of geographical location and task classification".

As a consequence, all tasks that were classified by the user in the same way and were carried out in the same geographical location, were assigned to the same "Context". Following the observations related to the user-activity distribution and observation of its bi-modality, the task classification used by the users in the diary study was based on one of the key factors affecting user performance. Users were asked in the study to classify the task as being either related to: work/university, pleasure or personal activity.

Factor	For Work Tasks	For Pleasure	For Personal Duties	Per task scope
Stress Level	37.48%	36.85%	66.67%	47.00%
Importance	11.11%	36.21%	20.41%	22.58%
Perceived difficulty	55.38%	56.05%	62.36%	57.93%
Estimated effort	55.90%	47.48%	49.49%	50.96%
Search query formulation difficulty	66.18%	65.85%	34.40%	55.48%
Perceived domain knowledge	64.29%	51.02%	36.36%	50.56%
Urgency	29.12%	52.81%	34.64%	38.86%
Confidence level	21.76%	31.25%	26.35%	26.45%

Table 6.2: The Coefficient of Variation of different moderating-factors measured for different task types and overall for the task classification used

Additionally the classification by Kellar et al. (2007), was used to observe the impact of finer granularity of task classification on behavioural variation. The classification is based on 4 distinct task types: fact-finding, information-gathering, browsing, and transactions.

Table 6.2 and Table 6.3 present the results of the data analysis. The results show that even when using the oversimplified definition of context (exclusively based on user, simple task classification and location) the variation of the moderating-factors has decreased. The biggest differences (Table 6.3) observed were the overall stress level associated to the particular context and the difficulty of formulating the search query. Overall, across the entire population, the stress level had a coefficient of variation of 72.58%. Just by grouping the data by the "Context" factor defined above, this coefficient dropped to 10.65%. Further introduction of the task classification based on Kellar et al. (2007) decreased the coefficient of variation to 7.1%.

Overall, the contextualisation of the experiment by the stratification of data, based on user location, and two types of task classification, led to a decrease of variability of the moderating-factors, as expected by the model. The decrease of moderating-factor variability was high in the case of all of the moderating-factors demonstrating the stability of user environments. The above results illustrate the stability of the user state in consistent environments.

Factor	Per Task Type in Context	Per Context	Per Task Type	Overall
Stress Level	**7.10%**	10.65%	66.07%	72.58%
Importance	**11.34%**	17.28%	25.06%	32.24%
Perceived difficulty	**11.78%**	24.43%	51.50%	57.65%
Estimated effort	**13.23%**	21.07%	35.46%	48.59%
Search query formulation difficulty	**15.53%**	22.48%	54.34%	62.71%
Perceived domain knowledge	**19.01%**	24.34%	49.10%	52.84%
Urgency	**16.39%**	24.59%	39.83%	49.38%
Confidence level	**7.98%**	11.44%	32.55%	30.19%

Table 6.3: The Coefficient of Variation of different moderating-factors for task type classification based on Kellar et al. (2007)

This is very consistent with the observation made in the previous section of this chapter, pointing out the logical nature of the process of work place and work time selection. The above supported the observation that even though people do make decisions based on a number of factors which are difficult to quantify, in similar situations, they are making consistent and logical decisions following the same reasoning process. This is consistent with the assumptions of the proposed approach, as well as the assumptions of ethnomethodological approaches. The application of the ethnomethodological method for IR system development is then feasible, as long as one can justify the stability of the key moderating-factors responsible for behavioural variance (e.g. what should be possible using the proposed methodology).

6.2.1.4 Task Frequency Distribution

Participants of User Study 1 reported a limited and fixed amount of problems that were very frequently reappearing in their everyday life. During the diary study the participants were asked to specify when they performed a particular reported task, how frequently this task was performed, and finally when, approximately, did they carry out this task for the first time. The data from the interview was not used in this analysis, as it was self-selected, and only the most repeatable tasks were reported by the participants, skewing the results in favour of repeatable interactions.

The data from the diary study was processed and all the tasks were grouped into three categories as illustrated in Figure 6.7. The results show that more than 49% of all interactions reported in the diary study were repeatable, which meant that participants declared that they had performed them frequently throughout the previous weeks and months.

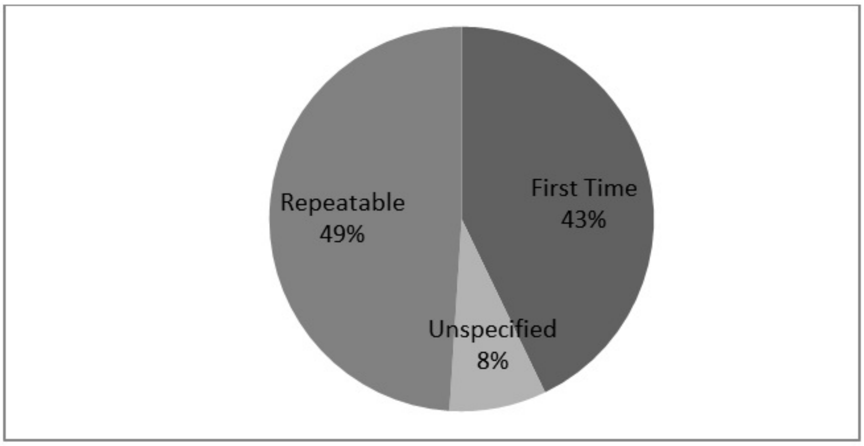


Figure 6.7: The classification of diary entries based on their repeatability -

Those observations were consistent with the general psychological research of habit development and automatic behaviour development. The results were also similar to the results reported by Wood et al. (2002). They had observed that not only were about 45% of reported activities highly repeatable, but also they were repeated in the same contexts and locations, which is consistent with the observations reported in this research.

6.2.1.5 Consistency of Repeatable Tasks

User Study 1 required the user to describe the task that is carried out, but also to assess whether it was a recurring task and if so when was the first time and what was the frequency. This information was used to classify the tasks as either being habitual or not based on the definition of habituation by Ouellette and Wood (1998); Wood et al. (2002)).

Based on that classification a frequency distribution of habitual behaviours was created (Figure 6.8). The analysis of the histogram of frequencies of habitual behaviours resembles a Log-Normal (Galton's) distribution (Figure 6.8).

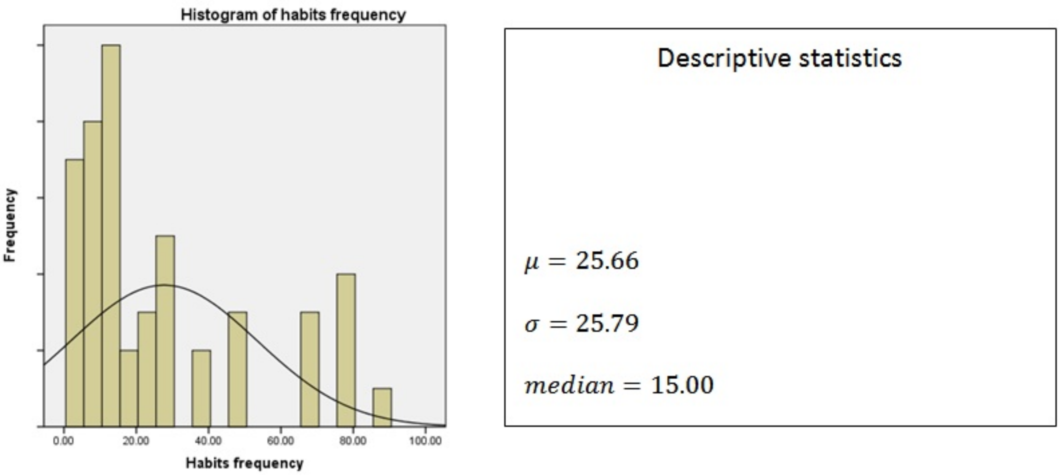


Figure 6.8: Habitual-behaviour frequency distribution function - Figure illustrates the log-normal nature of frequency distribution of habitual-behaviours

Further analysis using the descriptive statistics in a logarithmic domain confirms this initial observation (Figure 6.9).

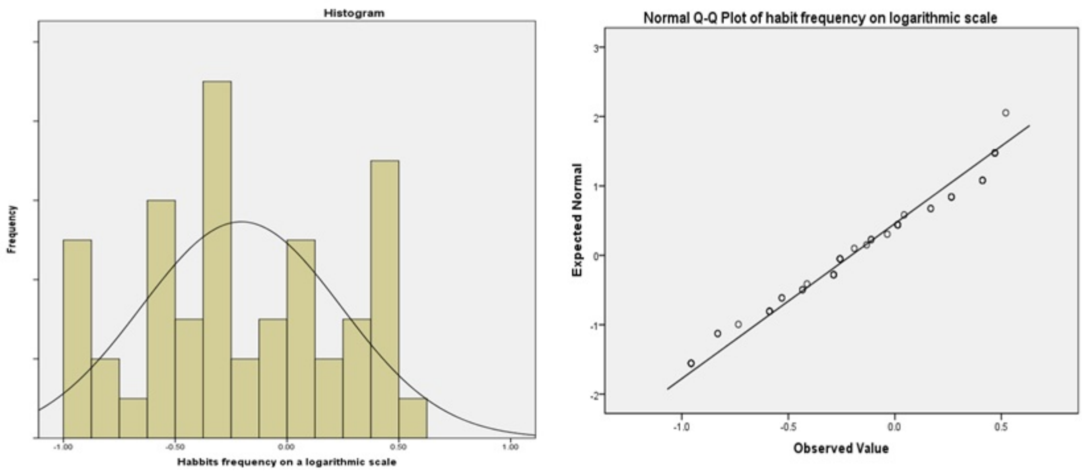


Figure 6.9: Habitual-behaviour frequency distribution function part 2 - Histogram of the habitual-behaviour frequency distribution function on a logarithmic scale

The Kolmogorov-Smirnov test of normality conducted in the logarithmic domain returned a very high value of p=.915 and so the assumption of normality (i.e. logarithmic)of the data cannot be rejected.

The analysis of the interaction frequency distribution, of both habitual and non-habitual behaviours on the exponential, scale is presented in Figure 6.10. One can observe a gap between the peak at 0 and the next peak at value 1. All of the interactions that fall into that region were classified in this study as single time / not repeatable interactions and are all equal to 0. The interactions between 1 and 15.2 were classified as non-habitual as they did not meet the criteria of habituation as outlined in 2002 by Wood et al. (i.e. they did not meet the requirements for stability).

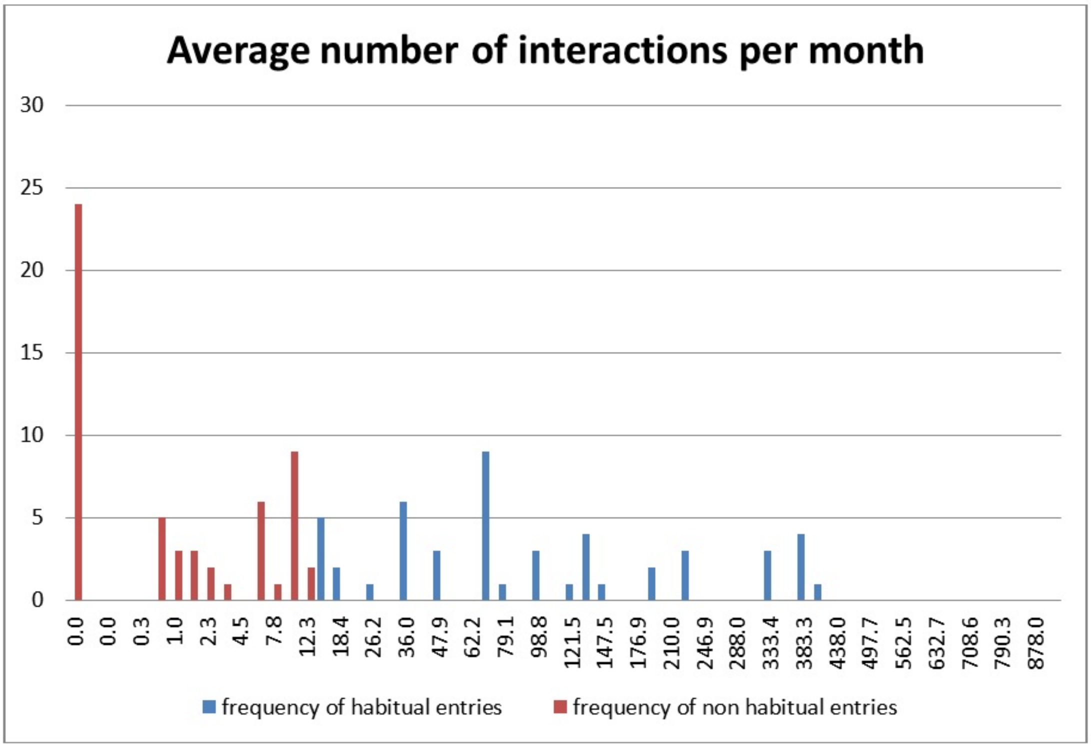


Figure 6.10: Distribution of interaction frequency on the exponential scale -

Figure 6.10 clearly illustrates the log normal distribution of interaction frequency. The results show that the Internet users developed a finite amount of frequent interactions that can be classified as habitual and that the quantity of interactions and their frequency was stable in time and followed a log-normal distribution.

In other words, and this was also supported by the qualitative information, the results of the study illustrate that the users engaged in a fairly small amount of very repeatable behaviours. The frequency of the secondary and tertiary tasks decreased exponentially. From the perspective of this research, and so from the perspective of the problem of consistency of HIB, those tasks were perfect candidates for further behavioural analysis as they were not only very frequent, but are distributed across a limited number of locations and are themselves very limited in numbers.

This observation is also very apparent in the qualitative data. User Study 1 users were very efficient in identifying and describing their primary tasks but had issues describing the secondary and tertiary tasks frequently explicitly and proactively dismissing their importance. The diary calendar view and the map view (see Appendices) was used to encourage participants to describe their sporadic work and enforce good task coverage. Participants did not feel comfortable describing those activities and when describing their work they were very clear about their unique, sporadic and abnormal nature.

A good example would be the analysis of the Internet use on mobile devices, which was encouraged at the end of the interview. When participants were asked about the goal, location and time distribution of this interaction few, if any of the participants could quantify it (note: the sample did not include any people working professionally with the mobile Internet) referring to the interaction as:

- being carried out "randomly, when there is a need for some information on the move"
- having: "almost no rules, maybe twice a month"
- happening: "almost never, every couple of days"

During the interviews, participants were able to recall exclusively the behaviours that, using the classification Wood et al. (2002), would be considered as habitual. Lower frequency activities were not described by the users and were captured only during the diary study.

What was also very interesting was the fact that frequently, sporadic and ad-hoc tasks were interrupted by the high frequency tasks and vice-versa. According to the participants only the highest concentration tasks or tasks carried out in very controlled environments were not interrupted. This has an implication on the personalised and contextualised IR system design as it blurs the notion of the current information need.

6.2.2 Habituation of Information Behaviour

Following the approach reported by Wood et al. (2002) the presence of habitual and automatic behaviour was investigated in the data captured in the diary studies. To allow for further comparison of the results of the current user study with the results obtained by Wood, Quinn, and Kashy the data was pre-processed in order to carry out the same statistical analysis.

It is important to notice that a big proportion of entries reported by the participants (i.e. 41.9%) met the outlined criteria for habituation. Moreover almost all of the interactions marked by the participants as repeatable met the higher criteria for habituation (i.e. 7.1% difference). This meant that interactions of high frequency were reported to also have a very high geographical stability and their execution rarely occurred in a different location. The average rating of context stability on a scale from 1 to 3 for the entire set of habitual behaviours was equal to M=1.98, SD=.34 which is a very high value, considering the richness of user information interaction.

Table 6.4 shows a comparison of our results with the results obtained by Wood et al. (2002).

The results are very similar to the original results reported by Wood et al. (2002) and support not only the hypothesis of user HIB consistency but also the hypothesis of the consistency of cognitive mechanisms responsible for IS&R and everyday activities.

6.2.3 Conclusions regarding the extent of HIB consistency

The goal of User Study 1 was to provide a foundation for further research and initially investigate the problem of HIB Consistency on a wider population of information users. The results of this study confirmed the initial research hypothesis. The results illustrated that a big proportion of participants' time was devoted to very repeatable IS&R tasks. More interestingly the results illustrated that those very repeatable and dominating tasks constituted only a small proportion of the overall number of tasks but were very consistent in terms of scope and associated user information behaviour. The remaining tasks identified in the study, even though varied in nature, location and time of interaction, were very infrequent or sporadic.

User Study 1 demonstrated that a majority of the frequent interactions were being carried out systematically in the same location, the same situational context and at

	This study		Wood,Quinn, & Kashy (2002) Study 1		Wood,Quinn, & Kashy (2002) Study 2	
Variable	M	SD	M	SD	M	SD
Number of diary entries per participant	12.25	3.095696	9.58	3.12	20.74	5.47
On the basis of the experimenter's rating, the proportion of behaviours classified as:						
habitual (performed almost daily, usually in same location)	0.419264	0.159396	0.35	0.19	0.43	0.16
corresponding with thoughts	0.478869	0.118158	0.61	0.19	0.53	0.16
On the basis of participants' ratings of each behaviour, the proportion of behaviours in which:						
other people were involved	0.141865	0.086531	0.49	0.18	0.44	0.16
Participants' ratings of :						
frequency of past performance	1.788877	0.363071	2.23	0.36	2.49	0.22
stability of context	1.98006	0.339423	2.55	0.34	2.57	0.23
intensity of emotions			1.86	0.42	1.82	0.31
attention required	2.535119	0.299584	N/A	N/A	2.27	0.38
behaviour difficulty	2.14881	0.597129	N/A	N/A	1.94	0.38
importance of behaviour for personal goals	3.526488	0.275382	N/A	N/A	2.47	0.63
amount of thought required before performance	1.591071	0.354136	N/A	N/A	2.19	0.56

Table 6.4: Comparison of general statistics from this user study with the results obtained by Wood et al. (2002)

similar times. As illustrated by the qualitative results the above was not a coincidence, but a result of situational, task and other constraints as well as user preferences. In all of the cases,, when asked, users were able to provide logical justification for their actions and recall a plan/rationale for the search process. Even though in some cases the decisions were made based on highly subjective and emotional factors (e.g. boring lecture) the responses to those factors were highly repeatable and frequently factored into the daily record. Examples of the above would include: taking a laptop selectively for "boring lectures", very high reported importance of communication with friends using Facebook, or the interruption of the primary task with side tasks (caused by anxiety).

What was also very interesting is the observation that across the repeatable tasks the distribution of other moderating-factors was very limited. The analysis of the coefficient of variance for other moderating-factors captured in the diary study clearly illustrates how stable emotionally, in terms of subjective difficulty, urgency, etc. were individual tasks. Using the moderating-factor vocabulary introduced in the previous chapters, the more defined the task the smaller is the distribution of the moderating-factors and as a consequence the smaller is the behavioural response observed.

The above does not mean that the moderating-factors are stable within longer interval of times. On the contrary, this study shows that users frequently interrupt the primary task or carry out multiple primary tasks at the same time. For example User 9 reported that when carrying out the university related work 40% of the overall time was spent on email, Facebook, Skype, various news portals, chatting or eating. This factor varied from participant to participant and was evenly distributed.

Finally the results of this research illustrated the consistency of work distribution with more generic psychological model of habitual behaviour and as such supported the hypothesis of consistency of the cognitive processes responsible for their control.

The results from User Study 1 provided justification for further research (User Studies 2 and 3). They illustrated the complexity of the problem of HIB consistency but also identified the fact that repeatable activities/tasks dominate the overall time spend on IR. More importantly the results identified only a very limited amount of those highly repeatable tasks on which further research can be focused. The above results are a justification for the ethnomethodological approach to HIB as well as IR research. The results suggested that for all of the users it should be possible to identify a very limited number of key primary tasks for which optimisation of IS&R utilities could be achieved. As such this is also an argument for the feasibility of contextualised and personalised IR.

6.3 Behavioural analysis

The diary study discussed in the previous section was followed by User Study 2 which aimed to assess the self-perceived user information behaviour consistency. User Study 2 did not focus on the generic IR population but on a population of Software Developers shortlisted for analysis in the previous stages of the research framework. User Study 2 was composed of an interview followed by a short questionnaire. It focused on the analysis of the behaviours identified in the previous step of the framework. The key focus was placed on:

- Opportunistic practices of developers.
- Reuse of small sections of code during IR.
- The way those information objects are retrieved and later reused in the development environment.

The study focused mainly on the analysis of the factors affecting the user decision to engage in the listed behaviours and the probability of those behaviours occurring frequently during normal development practice. The analysis was carried out following the Theory of Planned Behaviour framework and therefore the emphasis was placed on the analysis of:

- Users' attitude towards the identified behaviours
- Users' subjective norms
- Users' perception of behavioural control when planning/carrying out the behaviours

The rest of this section discusses the User Study 2 results and is structured around the above mentioned attitudes and perceptions.

Question	Actions per/day	Actions per/260 days
How often does your work related problems require from you the use of Internet/Intranet?	5.35	1391.91
How often do you look for code documentation and code examples?	2.12	550.00
How often do you reuse the code snippets of code from external sources to rework them in your projects?	0.56	146.41

Table 6.5: Frequency of Internet resource usage

6.3.1 Attitude towards behaviour

The attitude analysis was carried out to investigate the extent to which participants believed that their HIB will produce relevant information. The first stage of the analysis focuses on the perception of behaviour frequency which was treated in this study as a measure of attitude.

In order to measure their attitudes users were asked to rate them on a Likert scale. Table 6.5 shows the normalized frequencies of key behaviours on a ”per day” and ”per working year” basis. The values presented in the table below do not represent the number of search queries or search problems. They represent the number of work-related problems that were solved throughout the day. Each relating to numerous search iterations (as illustrated in User Study 3).

The results presented in Table 6.5 illustrated that the use of digital resources was critical for Software Developers' work activities. The interview information confirmed this observation.

Firstly, the user attitude towards various information objects/ information needs was listed in Figure 6.6. As can be noticed in Figure 6.6 only the search for code examples / snippets / modules was selected by all of the users. There was a high level of consistency in the answers provided by the users. During the interview, participants reported that they most frequently search for highly processed information (e.g. code examples / code solutions) in forums/blogs. Failure to find this information meant that they referred to more generic sources such as formal documentation.

What work related information do you look for on the internet?	Answer count per 16	Answer %
Code documentation.	11	68.75%
Code examples.	16	100.00%
Solutions to some code related problems.	14	87.50%
General ideas regarding the approach to take.	14	87.50%
Technological news.	7	43.75%
Technical material.	5	31.25%

Table 6.6: Frequency of Internet resource usage

Secondly, the interviews revealed that that code snippets/examples were the most valued form of information for the block level development problems. The free text information (with description of solutions and ideas) was preferred for subsystem level problems.

Thirdly, the way the information was searched and processed was analysed. The question attempted to identify what happens to the relevant information once it is found. Table 6.7 summarizes the most frequently reported actions. All of the participants reported that they either: copy-and-paste the relevant information or use it (e.g. by reading during development) to solve their problems.

Two responders reported that they often copied even loosely-related code, to obtain the template for further work. Also, two users reported that they frequently used the search engine if they did not remember the spelling of commands or groups of commands and used the retrieved content directly in their code. Overall the comments which related to the reuse of information were consistent with the research on opportunistic practices and psychology of programming (Brandt et al., 2009; Hartmann et al., 2008).

The above findings showed that the software developers being studied did not normally refer to abstract information such as documentation or platform architecture information. Instead, they preferred to retrieve partially relevant code and perform iterative experimentation on it. Moreover, the self-reported frequencies of interaction suggested that code reuse was not an unusual part of their activities but instead was

Once you find the information you need what do you usually do with it?	Answer count per 16	Answer %
Just read it and use the knowledge	11	68.75%
Copy the code snippet or code example	14	87.50%
Read the information in the same time solving the problem	13	81.25%
Print the information	3	18.75%
Bookmark, email or save the information in any way	7	43.75%
Share the information with other people	8	50.00%

Table 6.7: The reported usage of relevant information

very tightly coupled with their work.

The software developers' had a strong negative attitude towards the use of code-focused solutions such as Google Code or Source Forge and none of them reported their use in the study. All of the participants strongly preferred the use of general purpose search engines.

6.3.2 Subjective norm

The question regarding the reuse of online information during the code development (opportunistic programming) raised the highest amount of comments because of its unclear social status. That is most probably because some of the participants did not perceive it as fully ethical and had a need to caveat the extent of their code reuse. When asked about their opportunistic practices the participants reported that they used the code examples "mainly for very small problems where the solution is clearly defined" so in other words to distribute their cognition. Other participants were much more goal focused and reported that they did "whatever was required" to achieve the development goal quickly and to be in the position to go to the next task. On the other hand, there were some participants that had doubts regarding the morality of the opportunistic practices and reported that they engaged in them "Only through Internal code" claiming that they reuse only the company owned components.

Overall, the comments from the study suggested that the social perception of opportunistic practices was highly related to the other factor discussed in this study, that is the level of perceived behavioural control. Users did not perceive reuse of names, short sections of code or highly reoccurring elements as non-ethical or socially inappropriate. Participants found it acceptable to copy-and-paste the names of methods and classes or reuse snippets ("if I cannot remember the spelling or have found a good example"). They also did not perceive as unethical the process of development by example ("Copy and modify to suit needs"). The Developers perceived code reuse only if the Intellectual Property or licensing was an issue. Similarly, if the quality of the copied subsystem or system level content was recognised to be low or unpredictable, and if the issue was not addressed this could be perceived as unacceptable and in such case the reuse of code would be discouraged. Developers were not concerned about the novelty of their work directly but clearly focused on the task at hand. Hence the only social pressure related to their opportunistic practices was related primarely to licensing or quality issues. Developers perceived the reuse of open source code as socially appropriate and were happy to reuse "other people solutions to problems" giving them credit through the code in-line comments.

6.3.3 Perceived behavioural control

The participants in the study noted that their need to search for code related information in the Internet was highly driven by their current role in the project. This observation was very consistent with the results obtained by Freund et al. (2005) discussing the observed multi-modality of developer activities. This observation recurred across participants and was best represented by the following comments regarding the frequency of this interaction:

- "Sometimes I use it (the Internet) throughout the day but there are days that I do not use it not at all."
- "Often, but quite random as well, I could not classify, sometimes something like 1 to 10 times in a month really, sometimes nearing 10 a day."
- "Highly dependent on the stage of the project. At the early stages the interaction is very intensive and most of the work is Internet based. In the latter stages of

the project the level of interaction drops dramatically sometimes to the extent when the Internet is not needed at all."

Comments from the study suggested that the perception of control over the code related search was moderated by the pressure exerted either by the project deadlines or by management. Similarly, the developer's role in the project, was reported to moderate the user perception of control. For example, those developers who were involved in the support or test activities reported that they invested less time in the search of code related information. The tasks at hand and the role of the developer changed dynamically throughout the experiment. As a consequence the figures reported in Table 6.5, Table 6.6 and Table 6.7 represent an averaged value, skewed by the fact that there were periods of time when participants did not use the Internet for code related purposes. This observation, along with the actual values of interaction frequencies, will be further verified in User Study 3 discussed in the next section.

6.4 Identification of moderating-factors

User Study 2 was followed by the automated User Study 3. The goal of User Study 3 was twofold:

- To reinvestigate the question of Software Developers' HIB.
- To investigate the range of factors affecting user behavioural consistency, especially the shortlisted behaviours.

The structure of this section reflects the goals listed above.

6.4.1 Software Developers' Information Behaviour Consistency

6.4.1.1 Location Consistency

The analysis of location variability for the data collected in User Study 3 was much more complex than the same analysis in User Study 1 due to difficulty in automatic disambiguation of interaction location. The difficulty was related to the lack of precise geographical location and a need to derive this location from IP address information. As a consequence it was not possible to identify the movement of users across locations in order to establish the places of interaction (as defined in User Study 1). It was not possible to distinguish any of the home related interaction variations, because in all of the cases, the same wireless router was used in the home environment, and as a consequence, the same IP address was captured. The analysis was limited to the more generic location analysis. The analysis was made more complex by two additional properties of the IP address allocation being its often dynamic nature as well as the use of NAT technology. This occurred, especially in the company environment, which made differentiation of users possible only after merging the IP data with the user ID information.

The issue with the dynamic IP address was resolved by using one of the freely available IP Finder applications. Once the application received the enquired IP address it provided the name of the provider and the approximate geographical location of the experiment participant. Overall more than 300 unique IP addresses were retrieved from the IIS logs using the Microsoft IIS Log Parser. After removal of the automated traffic (generated by various Internet robots) and removal of the random traffic of non-authenticated users, 72 distinct IP addresses were identified. The IP addresses were processed manually to identify 15 unique interaction locations. The reported figure was very consistent with the information reported in the diary study for the same number of users. A value lower than in the diary was expected, owing to the inability to install the observation software (perhaps due to lack of support,, security concerns or lack of will.) on all of the platforms used referred to in the diaries.

Based on the data from the three different experimental tools (interview, diary and the automated observation), the graph showing distribution of frequency of location selection was created. The analysis presented below was based on the analysis of information regarding interaction location as well as the reported frequency interaction. The information in the figure below varies according to the research tools used across the three user studies.

- Interview: During the interview participants were explicitly asked about the location they used and the frequency of interactions at that location. The locations with frequency equal to 0 represented random interactions which according to the participants were highly unlikely to reoccur. The values represent the number of times the interaction took place per week.

- Diary data: Users were requested to record the location of the interaction as well as the frequency with which the interaction occurred. The values in the graph below represent a sum of those self-reported figures across the particular location, and so as a consequence, give the geographical frequency distribution of information interaction sessions.

- Log data: The content of the IIS server logs was analysed in order to extract the geographical location of the users as described in the procedure above. This information was merged with the reporting database and every reported page was assigned a unique geographical location. Together with the known time distribution of interaction this information was used to generate the data below which represents number of visited pages in a particular location per week.

Figure 6.11 illustrates the results of this analysis. The goal of this analysis was to illustrate the very high level of stability of the information interaction location selection in the study.

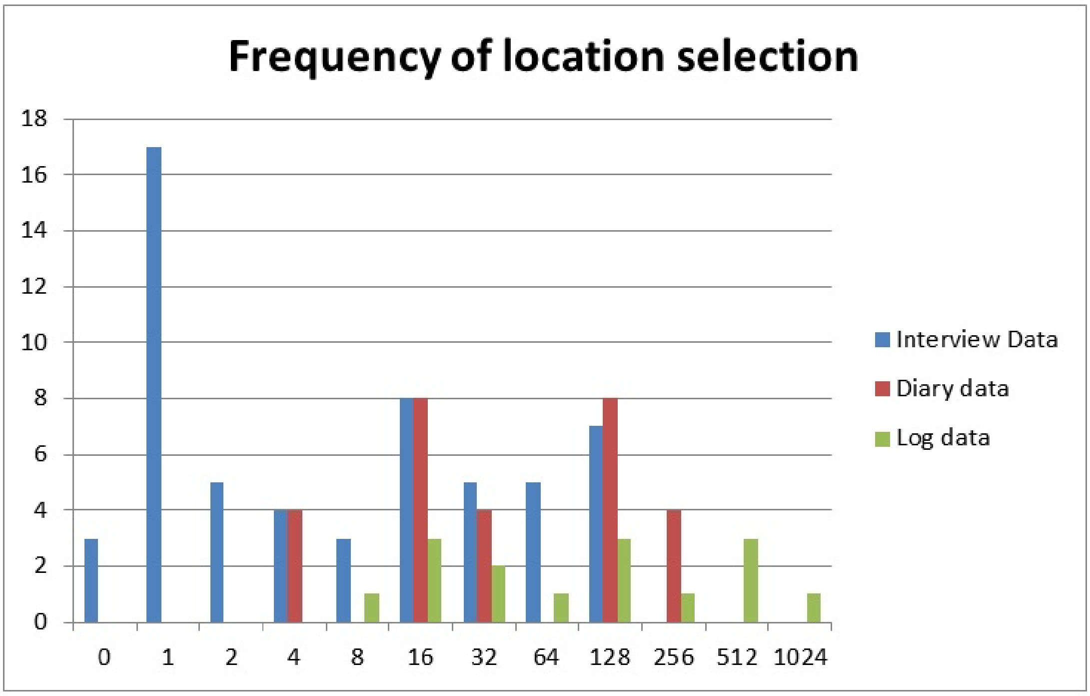


Figure 6.11: Frequency of location selection -

It is important to notice the number of locations which are selected infrequently by participants is proportional to the number of locations that are selected frequently. As expected, because of the differences between the scales used in all three studies (frequency of location selection, frequency of the sessions and frequency of page visits) the frequency of interactions from the diary information is shifted to the right in comparison to the frequency of interactions from the interview data and the frequency of interactions from the log data is shifted even more to the right.

6.4.1.2 Time Consistency

The results of User Study 1 were used to form the initial observations regarding the probability distribution of user activity (type and parameters) and the variance within and between the groups being studied. Those initial observations were then further verified by a large scale automated observation of user behaviour carried out in User Study 3. User study 3 focuses on the same aspects of user interaction. It does not rely on self-reported data which is limited in terms of quality and quantity but on objective interaction observation, therefore the distribution information is very accurate. Also due to the amount of data captured in User Study 3 it was possible to investigate the variation of data within and between the groups as well, with a higher level of accuracy than with the other methods.

Again it is crucial to note that the interpretation of the data from User Study 3 was different due to the different variables used. User study 3 based the analysis on observation of:

- User behaviour assosiated to individual web page visits.
- User behaviour accross multiple pages contained in a single "report file". This level of granularity was used where the IIS (web server used) log inforamtion was required.

This made the data different from the self-reported diary entries and calendar reports captured in User Study 1. The data captured in User Study 3 was more accurate in a sense that it captured user interaction with much greater atomicity than User Study 1 and was objective rather than self-reported. On the other hand not all of the users' environments could be monitored by the experimental software, as some of the environments where the interaction took place were not under the control of the user (e.g. work/university environments).

The analysis of User Study 3 started with the exploration of data and tests of normality. The data analysis (Figure 6.12) confirmed the initial observation, based on histograms and box plots regarding a non-normal nature of user activity time distribution.

Tests of Normality						
	Kolmogorov-Smirnova			Shapiro-Wilk		
	Statistic	df	Sig.	Statistic	df	Sig.
User10	.159	2158	.000	.946	2158	.000
User18	.110	4057	.000	.956	4057	.000
User20	.182	1806	.000	.924	1806	.000
User21	.101	808	.000	.947	808	.000
User22	.136	3941	.000	.955	3941	.000
User23	.100	1913	.000	.973	1913	.000
User24	.094	3843	.000	.975	3843	.000
User26	.109	1837	.000	.951	1837	.000
User28	.123	1456	.000	.959	1456	.000
User29	.071	1827	.000	.966	1827	.000
User30	.138	691	.000	.943	691	.000
User32	.133	430	.000	.932	430	.000

Figure 6.12: Tests of normality of user data -

The analysis of the User Study 3 data confirmed the initial observation suggesting that the time distribution followed a mixed (binomial) model. Further transformations of the data did not allow for its further normalisation. As a consequence non-parametric tests were used for data analysis.

Figure 6.13 presents the weekly distribution of activities of an exemplary user. Various colours represent different locations of interaction and once again suggest a mixed-model of activity time distribution.

A closer analysis of user 22 daily activity distributions across time and location confirmed the observation of User Study 1. Figure 6.14 and 6.15 illustrates the impact of the location, associated tasks and context on the user daily routine. The activity

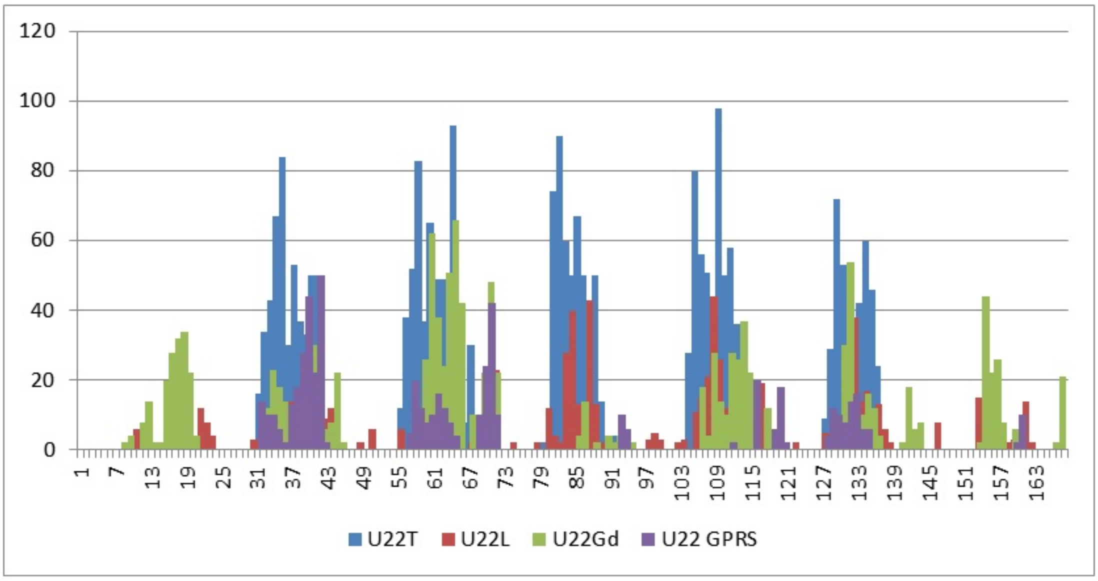


Figure 6.13: User activity time distribution - IIS logs based User 22 activity time based distribution. The graph represents frequency of experimental web service calls therefore in practice represents more than one page interaction report.

across different locations varies and all of the parameters of the distribution function are observed to change. This includes both the mean and median frequency of interaction time as well as the magnitude of the frequency distribution itself.

Overall (Figure 6.16 and Figure 6.17) the activity distribution of user 22 indeed proves to be a complex mixed-model with at least 4 distinct random variables (not necessarily independent). The only trend exposed by the data is the division of distribution into two key clusters and its bimodal nature. This observation further reinforces the role task and interaction time plays in HIB.

Further analysis of activities of other users confirms the above discussed observations (Figure 17). The expectation-maximisation analysis of distribution again provides a mixed-model ($\omega_0 = .73, \mu_0 = 14.09, \sigma_0^2 = 5.70, \omega_1 = .22, \mu_1 = 21.62, \sigma_1^2 = 2.81$) with a statistically significant difference between the normal distributions of the model (z=3.20,K-S D=.45,p¡.01).

The individual differences between the users are significant (more discussion regarding variation analysis will be carried out in the following sections). The results illustrate the impact of personal factors on HIB and its consistency. Figure 6.19 illustrates the extent of the differences between the individual users. The differences between the users

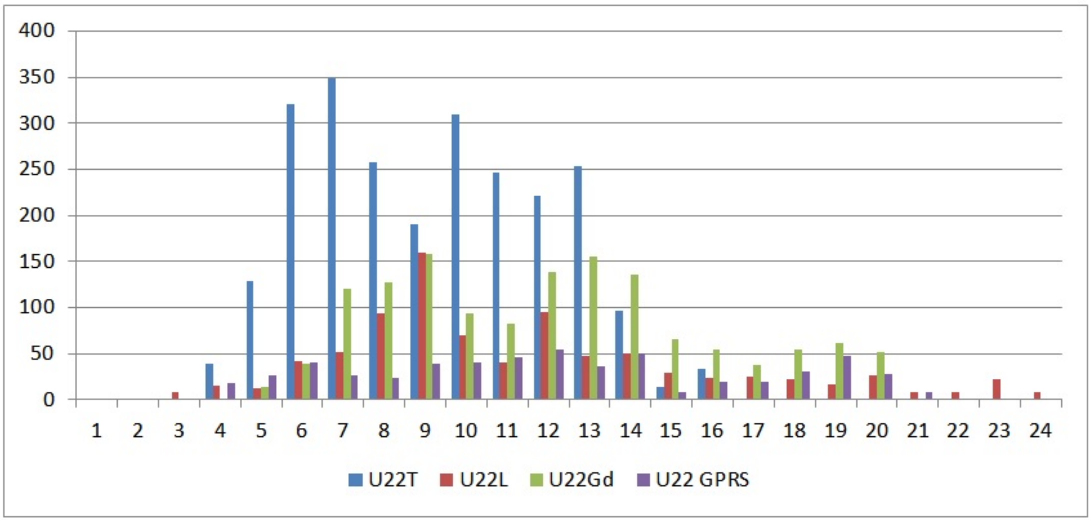


Figure 6.14: User 22 daily activity time distribution -

are statistically significant. User 22 ($\omega_0 = .73, \mu_0 = 14.09, \sigma_0^2 = 5.70, \omega_1 = .22, \mu_1 = 21.62, \sigma_1^2 = 2.81, z = 3.20, K - SD = .45, p < .01$) is more focused on Internet interaction during the first phase of interaction than the User 10 ($\omega_0 = .23, \mu_0 = 13.26, \sigma_0^2 = 2.58, \omega_1 = .77, \mu_1 = 20.11, \sigma_1^2 = 4.77, z = 3.55, K - SD = .50, p < .01$), the interaction starts earlier and the first phase dominates over the second phase thus marginalising its influence on the overall distribution.

In all 12 cases the above mentioned mixed-model composed of two normal distributions was observed. This is a reoccurring trend across all of the participants. Even though the analysis of activity distribution in different locations identified a number of other latent random variables they did not reoccur across the participants. The above is true also for the data gathered through the interviews as well as the diary study.

The analysis of the interview, diary and automated observation data illustrates a significant magnitude of differences between the study participants. Further analysis of variance will be carried out in the following section.

The results of normality tests presented in Table 6.8 clearly illustrate that the activity time distribution of participants in neither of the cases could be treated as normal. Exploration of data illustrates a very high likelihood of the bimodal nature of the user-activity distribution. Further application of expectation maximisation methods on the data confirmed the bimodal nature of the distributions. Table 6.8 illustrates the results of the users for which the Kolmogorov-Smirnoff test returned statistically significant

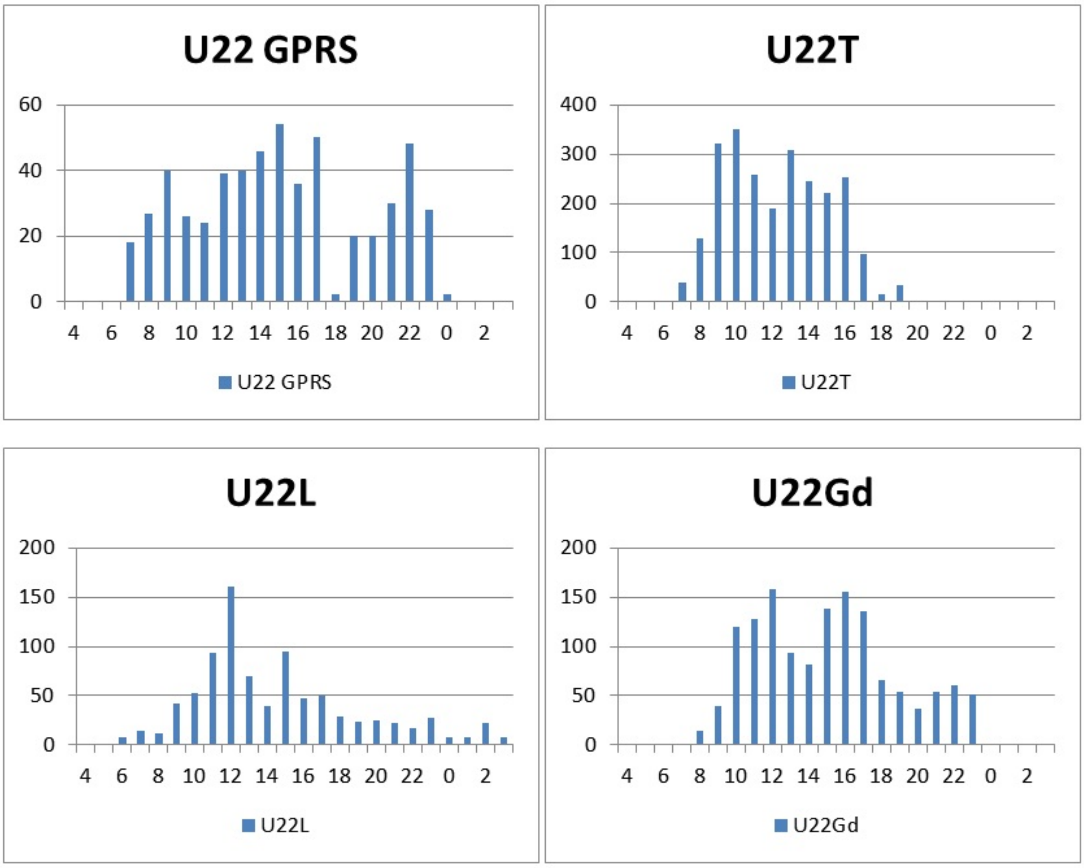


Figure 6.15: User 22 daily activity time distribution split out into different locations -

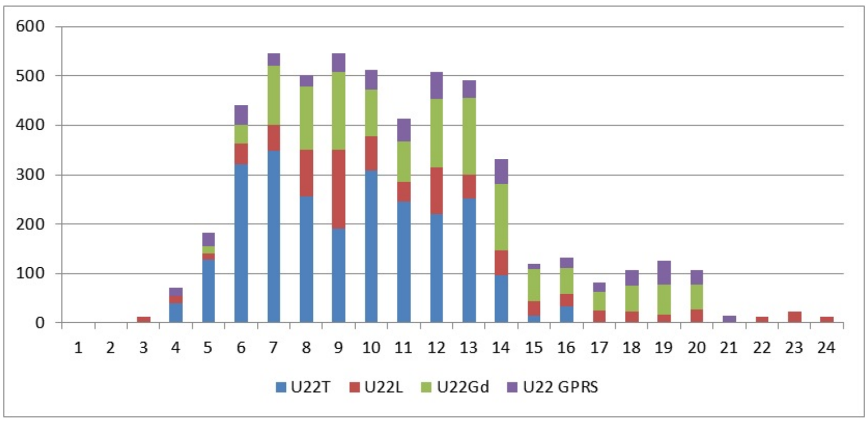


Figure 6.16: User 22 location and time distribution - The activity time distribution of user 22 HIB. Different colours in the histogram represent distinct locations reported in the study. Their size represents the observed frequency.

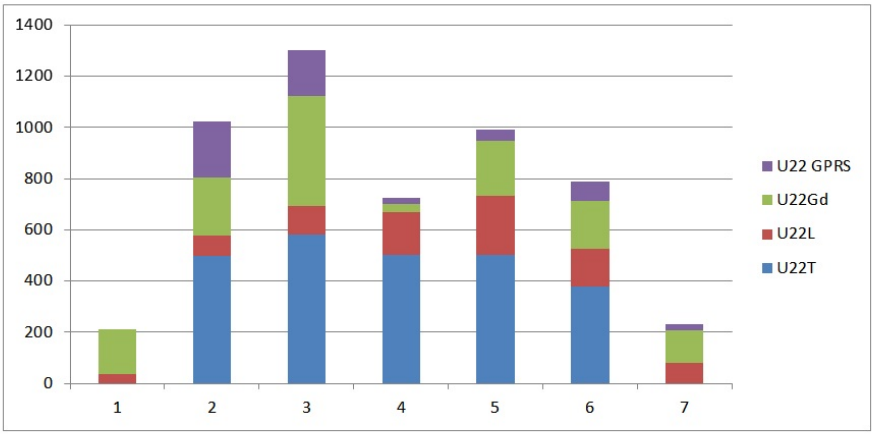
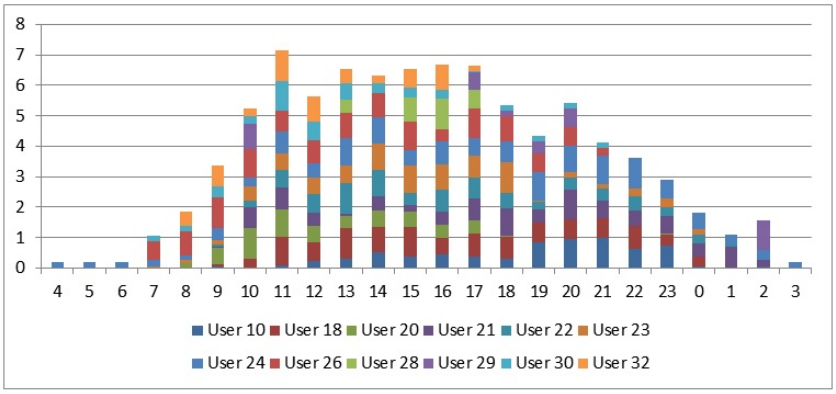


Figure 6.17: User 22 weekly location and time distribution - The activity weekly time-distribution of user 22 HIB. Different colours in the histogram represent distinct locations reported in the study. Their size represents the observed frequency.



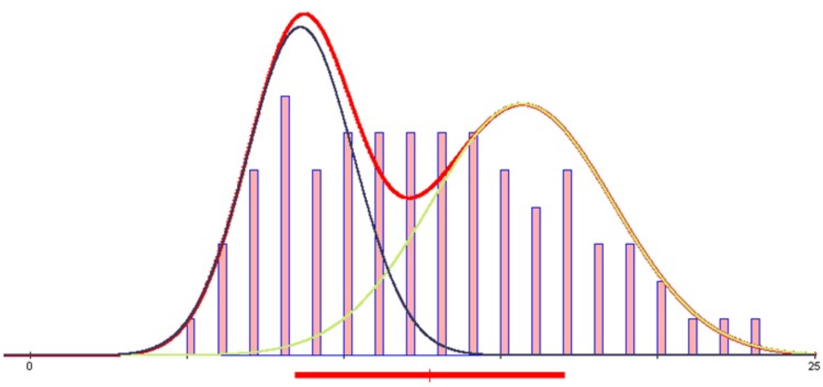


Figure 6.18: All user-activity distribution histogram and the histogram based mixed-model -

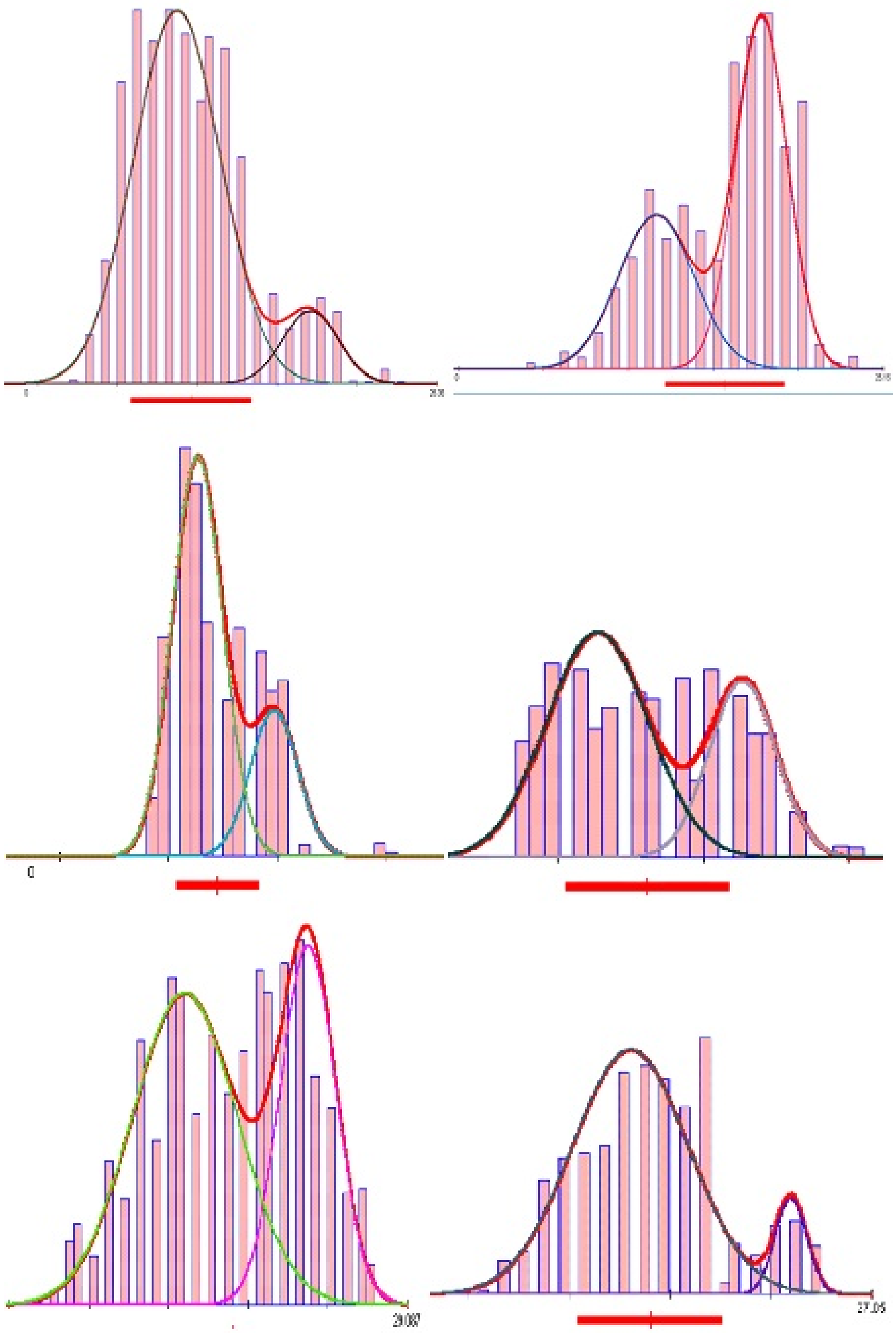


Figure 6.19: User Activity Distribution - Comparison of activity distribution function of User 22 (upper left) and User 10 (upper right), User 20 (middle left), User 26 (middle right), User 24 (lower left) and User 23 (lower right)

	User 10	User 18	User 20	User 21	User 22	User 23	User 24	User 26	User 30	User 32
$\omega 0$	0.23	0.8	0.74	0.26	0.73	0.9	0.62	0.66	0.8	0.67
$\mu 0$	13.26	14.99	11.12	11.49	14.09	14.62	13.98	11.12	11.26	10.66
$\sigma^2 0$	2.58	8.85	2.81	2.13	5.7	9	12.74	6.46	2.88	2.52
$\omega 1$	0.77	0.2	0.25	0.74	0.22	0.1	0.38	0.34	0.2	0.33
$\mu 1$	20.11	22.15	16.25	20.37	21.62	22.92	22.38	18.39	16.49	15.42
$\sigma^2 1$	4.77	1.03	2.52	9.27	2.81	0.65	3.63	2.84	3.59	0.68
z	3.55	4.04	2.99	3.34	3.2	5.59	1.72	2.21	3.62	2.42
K-S D	0.5	0.57	0.42	0.47	0.45	0.79	0.24	0.31	0.51	0.34
p	<.01	<.01	<.01	<.01	<.01	<.01	<.01	<.01	<.01	<.01

Table 6.8: Result of expectation maximisation algorithm applied to the data.

difference between models 0 and 1. For users 28 and 29 results were not significant at a level of p=.01.

Because of the shape of the observed time distribution (mixed-model resembling bimodal distribution) the parametric tests of variance (ANOVA) could not be safely carried out.

Secondly the tests of homoscedasticity (homogeneity of variance), carried out visually (Figure 6.20) and statistically using the Levene's test (F(11,21504)=177.64,p=.000), illustrate that there exist big differences in user interaction-time variance. This observation was expected due to hypothesised individual differences in user interaction. This indicates a difference between the groups (the Kruskal-Wallis test assumes homogeneity of variance and as consequence cannot be safely used).

The next analysis focuses on the coefficient of variability and the extent to which it changes with the introduction of the two models obtained by the expectation maximisation algorithm. Table 6.9 illustrates a decrease of the value of the coefficient of variance which is very similar to the decrease observed when task granularity is being introduced further in this chapter. This table illustrates the impact of the work and personal-activity differentiation which is very often discussed in IR literature (Case, 2007).

What is important to note is the observation that the coefficient is not distributed evenly across the experiment participants and that the particular pattern of interaction and the extent to which this bimodality influences user interaction is very specific for

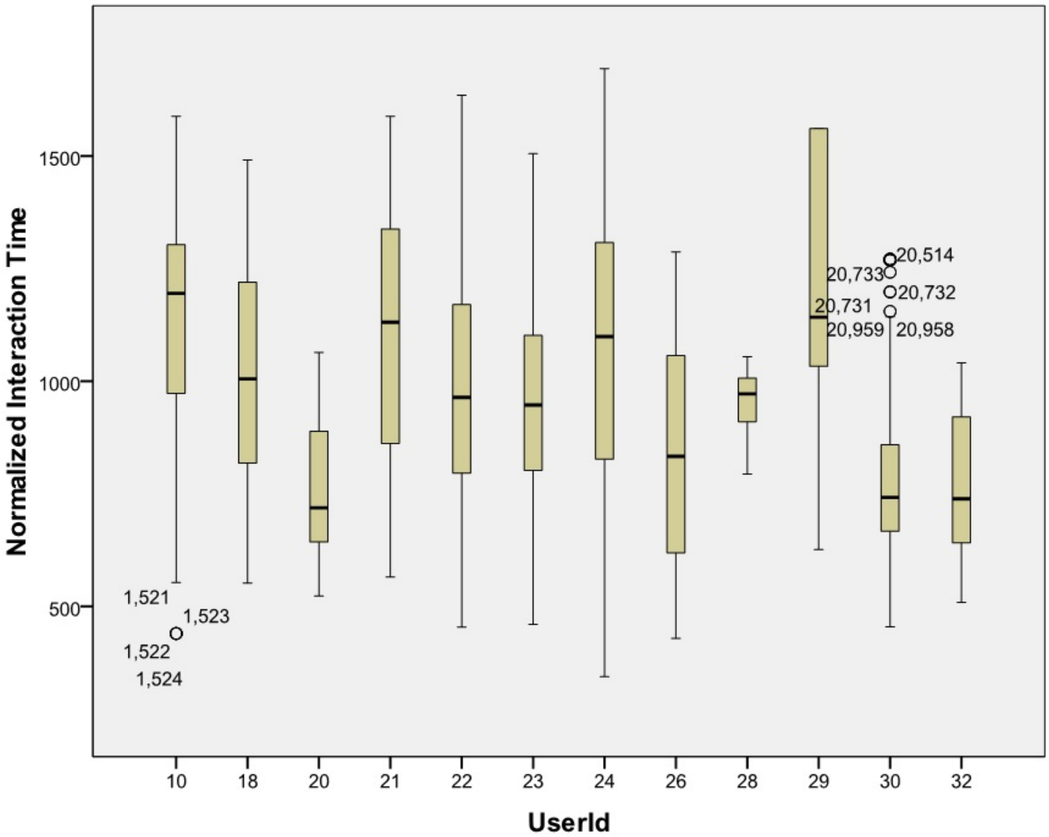


Figure 6.20: Interaction-time box plot - The Box Plot illustrating user interaction time. Numbers on the bottom represent user IDs

	User 10	User 18	User 20	User 21	User 22	User 23	User 24	User 26	User 30	User 32
Model 0	12.11%	19.85%	15.07%	12.70%	16.94%	20.52%	25.53%	22.86%	15.07%	14.89%
Model 1	10.86%	4.58%	9.77%	14.95%	7.75%	3.52%	8.51%	9.16%	11.49%	5.35%
Bimodal	19.10%	24.00%	21.02%	26.41%	24.86%	24.41%	29.59%	29.85%	22.15%	21.57%

Table 6.9: Analysis of coefficients of variability of User Study 3 users for bimodal distribution and two extracted normal models

a particular user in a particular time. Finally the results illustrate clearly a variation of information environment selection between the users and rather small variation for the users themselves supporting the hypothesis of the environmental consistency.

6.4.1.3 Task Distribution Consistency

A similar analysis was carried out based on the data collected in User Study 3. User study 3 did not collect the data regarding the actual task directly therefore a substitute variable had to be generated based on the available data. User study 3 collected the URL of the pages that were visited along with the interaction time and the ID of the user that was interacting with the information resource. This data was used to generate the frequency plot illustrated in Figure 6.21. The plot was generated through the extraction of the domain names of the pages visited and calculating the frequency of visits on a per user basis.

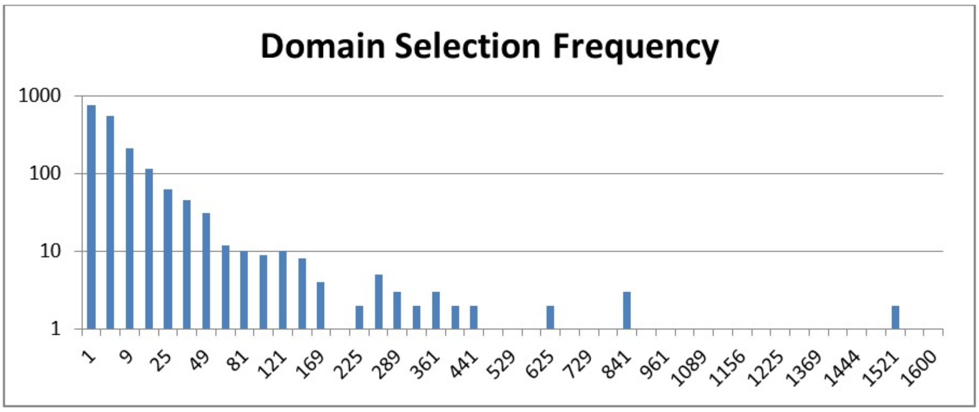


Figure 6.21: Domain-selection frequency on a logarithmic scale based on User Study 3 -

In the User Study 3 the user interaction was automatically captured. Therefore the quantity of the observed page visits is higher than the amount of data captured in User Study 1 which was diary based. This allows for more detailed analysis of user task distribution and illustration of the real extent of user task repeatability. As part of the data analysis each page interaction was classified into one of the three groups depending on the observed frequency of visits:

1. First time interaction: the page was classified to this group if a particular user interacted with the domain to which it belongs only once through the course of the study;

2. Infrequent interaction: the page was classified to this group if a particular user interacted with the domain to which the page belongs N or less times where N is an amendable threshold value.

3. Frequent interaction: the page was classified to this group if the number of domain visits was higher than the threshold value.

The goal of this analysis was to illustrate the proportion of the three classes of interaction depending on the threshold value. Figure 6.22 illustrates the results of analysis with the threshold value changing between 30 domain visits per user per experiment period to 120 visits per user per experiment period. The results show that only 4% of page visits were addressed at the unique domains. Moreover the data illustrates that even with a very high threshold value of 120 domain visits more than 40% of interactions were highly repeatable. This result is also consistent with the observations from the psychological research on habit development and overall repeatability of human behaviour (Wood et al., 2002).

The above analysis shows that a big proportion of information interactions are highly repeatable in terms of the task domain.

6.4.1.4 Consistency of Repeatable Tasks

The consistency of repeatable tasks was analysed in a same way as in User Study 1 this time on data captured automatically through User Study 3. Since the task information was not directly available (as users were not asked to provide it during the interaction) the task had to be inferred from the observed data. In order to carry out the above mentioned analysis, the interactions of the user were grouped around the network domains of the sites visited. This approach to task identification was very efficient as far as the automated analysis was concerned as it can be semi-automated. The tasks, identified using this approach, are different to the tasks captured in User Study 1. The tasks classified based on the domain analysis are of the lowest possible

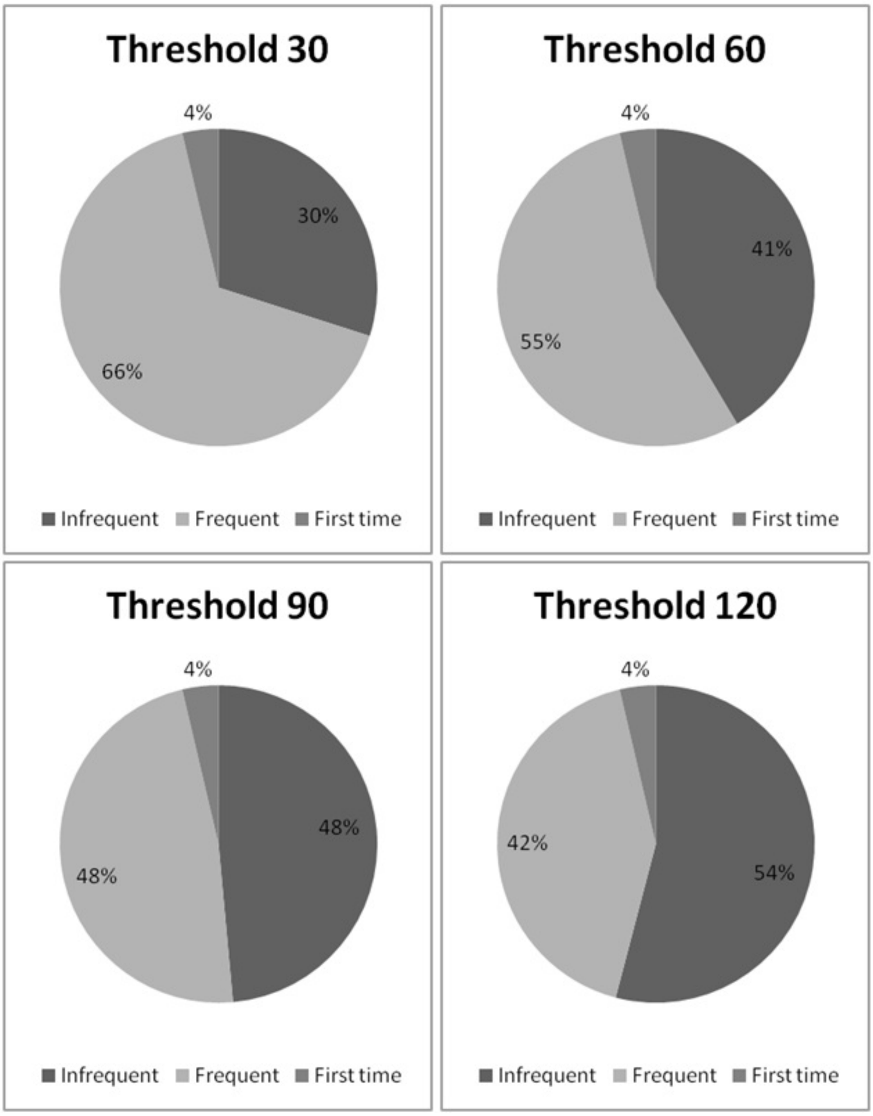


Figure 6.22: Repeatability of activities - The classification of log entries based on the user frequency of visits. Charts represent distinct visit classification depending on the classification threshold.

granularity in contrast to tasks in User Study 3 which were much more abstract. This does not affect the overall data analysis or the data frequency distribution.

Further analysis focuses on the investigation of visit frequency distribution and its variance. The analysis of the histograms of the frequency distribution presented in Figure 6.23 suggests a non-normal distribution of the data. The histogram resembles more closely the exponential or log normal distribution.

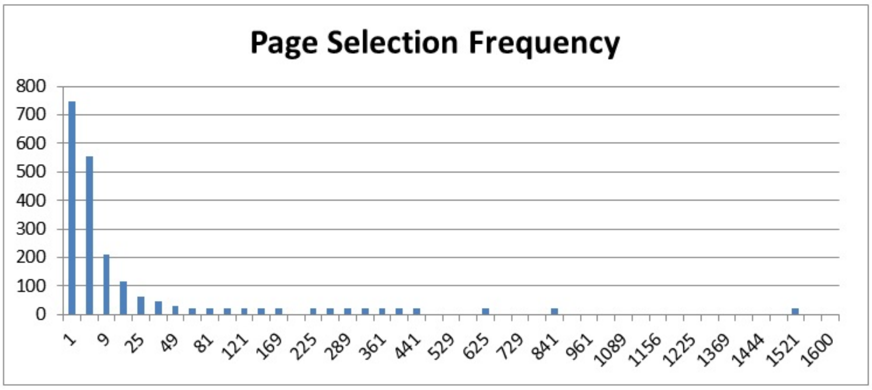


Figure 6.23: Page selection frequency for all users -

The Kolmogorov-Smirnov test of normality failed to support the hypothesis of normality as well as the hypothesis of the exponential distribution for all 12 users (in both cases all p¡ .000).

The results of User Study 1 supported the hypothesis of log-normal nature of this distribution. Therefore it is expected that the log-normal distribution will also be observed when analysing a much bigger data set. Figure 6.24 illustrates a histogram of normalised frequency data in a logarithmic scale which reveals the log-normal nature of the data. Because of the differences in the experimental design, frequent and infrequent interaction could be differentiated exclusively, based on the user activity within the experiment time. The fact that the distribution of data was based on a finite time interval meant that there will be a lower bound of data observed which is unique to a participants' experiment duration and equal to $1/(numberofweeks)$. The interactions that occurred with this particular frequency or with a lower frequency were not distinguishable in our experiment and manifested themselves as a number of peaks on the left hand side of the logarithmic histogram.

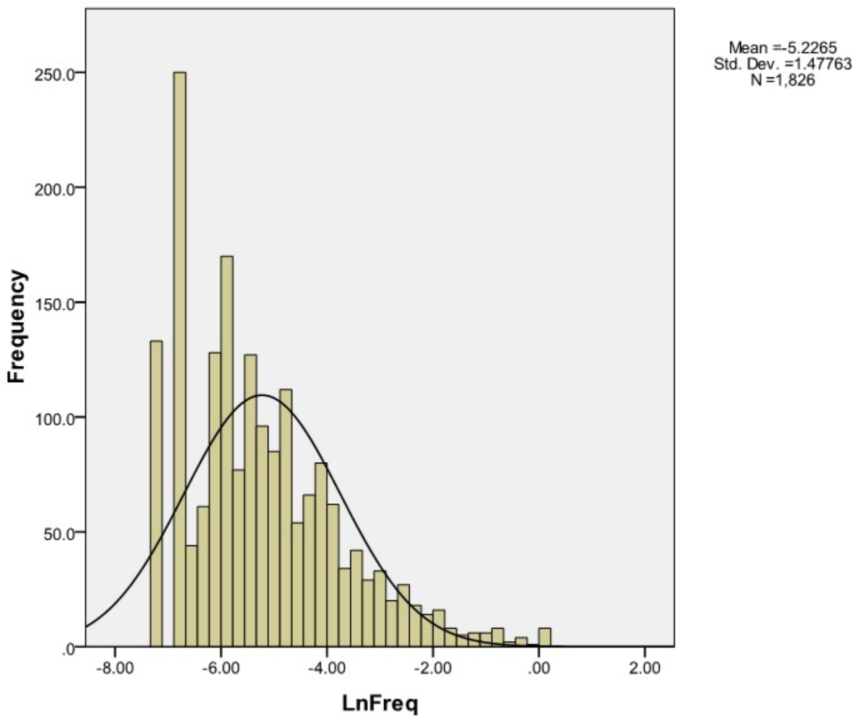


Figure 6.24: Histogram of user-activity distribution in logarithmic scale -

Following the above observation the expectation maximisation algorithm was employed to establish the parameters of the log-normal distribution based on the captured data (Figure 6.25). The generation of the model was followed by the Kolmogorov-Smirnoff which with p=.60 did not allow the rejection of the null hypothesis regarding the similarity of the data and the generated model.

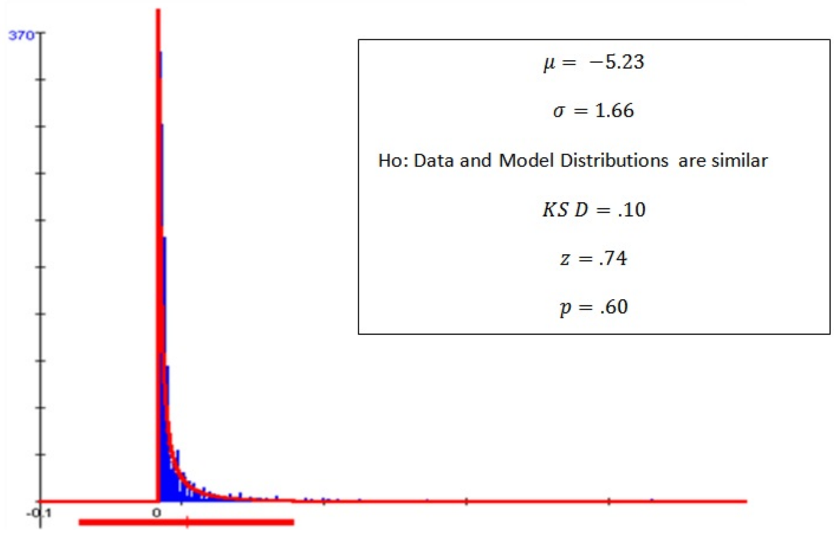


Figure 6.25: Log normal distribution based on expectation maximisation algorithm -

The Levene's test of homogeneity of variance confirmed the assumption of distribution homogeneity for all of the users (F(11,1837)= .860,p=.579). As a consequence the Kruskal-Walis non parametric test of variance could be carried out. The test indicated that the null hypothesis regarding the equality of the frequency distribution populations ($\chi^2(11, N = 1837) = 16.263, p = .132$) could not be rejected. The differences between the users were not statistically significant indicating the consistency of user task frequency distribution.

The above analysis does not take into account the geographical distribution of user activity. In order to address the above the log information containing IP addresses of the users was merged with the detailed information regarding user activity. The IP information was pre-processed to remove the impact of the dynamic IP which is supplied by the majority of ISP. The IP information was processed manually using the "who is" service in combination with the Machine Id information to produce unique user location IDs. The data regarding user activity was then processed to extract the domains of the pages visited which were further used as a base for task analysis as discussed below

Figure 6.26 illustrates the results of the task-location analysis. The figure illustrates that a majority of interactions were carried out in a single location (80.43%). Only one task of a single user was reported to be carried out in more than five locations therefore it was treated as an outlier and not included in further analysis.

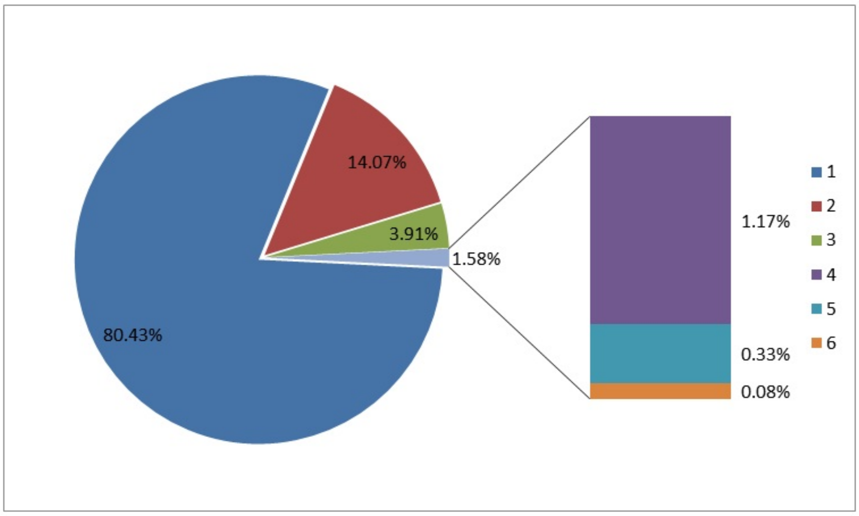


Figure 6.26: Task distribution across different locations -

In order to investigate the distribution of tasks across different locations the tasks were grouped based on the number of locations they were observed in. The grouping was further used for the frequency distribution analysis which revealed the types of tasks, one off/frequent/infrequent, carried out across locations. Figure 6.27 and Figure 6.28 illustrate the results of this analysis and suggest differences between the different populations.

Once again the frequency data has a log normal distribution with a left hand distortion related to the time constraints of the study (note: this effect was discussed in the above analysis). The analysis of the results illustrates that the interactions which were carried out in just one or two distinct locations span the entire frequency spectrum but have a logarithmic mean at around 0. The tasks which were carried out in more than two locations were shifted to the right, so towards much higher frequencies.

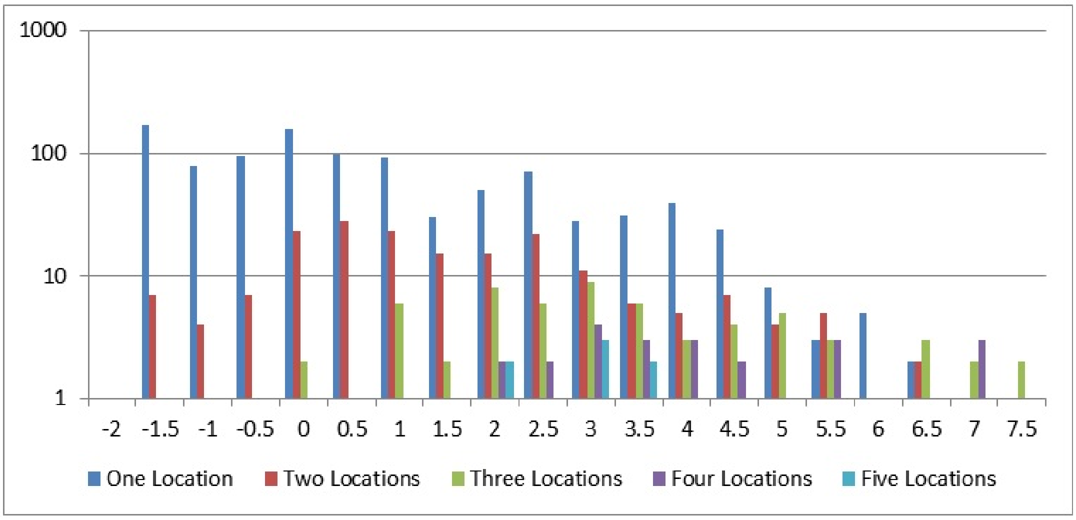


Figure 6.27: Natural logarithm task frequency distribution across different locations in logarithmic scale -

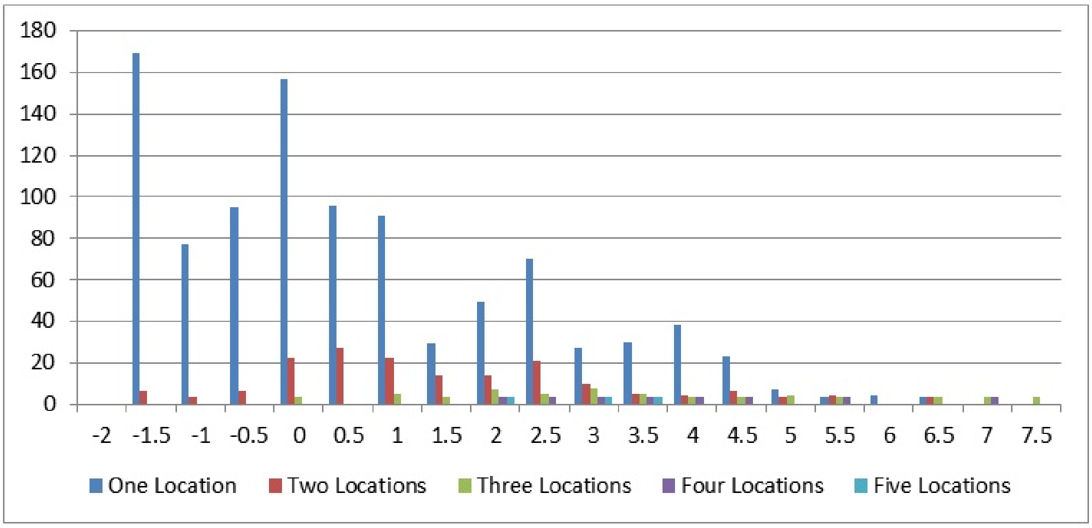


Figure 6.28: Natural logarithm task frequency distribution across different locations -

Number of Locations	1	2	3	4	5
1		$F(1{,}1131) = .026, p = .871$		$F(1{,}1197) = 65.484, p = .000$	
2					
3					
4				$F(2{,}62) = .916, p = .416$	
5					

Table 6.10: Levene's test of homogeneity of variance for task carried out in different number of locations

Interactions that took place in multiple locations were very infrequent in comparison with the remaining activities.

The above is consistent with the statistical analysis. The Levene's test (Table 6.10) failed to support the hypothesis of homogeneity of variance between the tasks carried out in one and two locations and the rest of tasks indicating statistically significant difference between those two groups. On the other hand the Levene's test supported the hypothesis of the homogeneity of variance between groups one and two as well as three, four and five.

The Kruska-Wallis non-parametric test indicated that the task frequency distributions for tasks carried out in one and two locations were different ($\chi^2(1, N = 1133) = 14.942, p = .000$). On the other hand the null hypothesis regarding the similarities of task frequency distributions for tasks carried in more than two places ($\chi^2(2, N = 62) = 2.656, p = .265$) cannot be rejected.

The above results illustrate clearly that not only were the distribution of tasks consistent for individual users but also that there were no statistically significant differences between the normalised distributions across different users. Moreover, the results illustrate that the tasks which were carried out in one location had a different frequency distribution than the rest of the tasks. Similarly the frequency of tasks which were carried out in two locations was different from that of the rest of the tasks although the variance was homogenous with the tasks that were carried out in one location. Finally, the frequency of the tasks that were carried out in more than two places was the same.

The above mentioned results confirm the hypothesis of task distribution consistency for User Study 3.

6.4.2 Identified behaviours and impact of moderation

The goal of this stage of the User Study 3 result analysis is to illustrate the impact of moderating-factors on the following shortlisted behaviours, including:

- Direct code reuse practices at block level (distribution of cognition)
- Direct code reuse at subsystem and system level (opportunistic programming)
- Direct reuse of system level components e.g. libraries, classes, frameworks and bigger components (component reuse)
- Indirect code and information reuse practices at different complexity levels (block, subsystem, system)

Because of the structure of the experimental software and as a consequence the structure of the experimental results the analysis focuses on the key information objects captured:

- Directly reused code elements including all three types of behaviour. The behaviours were further subdivided into three categories based on the overall copied content length and subjectively annotated complexity.
- Indirectly reused information.

6.4.2.1 Direct opportunistic practices: code reuse

A total of 1296 non-trivial copy-and-pastes (originating out of the report which lasted for more than 20 seconds) had been recorded, out of which 379 originated from outside of the development environment. The analysis revealed substantial behaviour differences between the interactions when the information was sourced from the browser and when it was sourced from other location. The results indicated that the presence of copy-and-pasting from the browser was an indicator of higher problem complexity, manifested through longer active times, captured in the behavioural reports, but also as far as the complexity of the copy-and-pasted content was concerned (measured through the content length and subjective complexity).

Figure 6.29 demonstrates how the active time of interaction with the source code document changes, depending on which sources were used for the purpose of copy-and-pasting. It is important to note the very strict definition of active time that was used in this study. The active time was the time the user spends actually interacting with the source code document. The active-time was calculated almost exclusively when the user was providing input to the development environment or interacting with it in other ways (for example building, debugging, etc.), it was not captured when there was no focus on the document or when the document was idle. As a consequence of being able to get the observation data directly from the development environment and creating such a strict definition the active time as measured here is a very good approximation of the amount of direct attention that was paid to the development document.

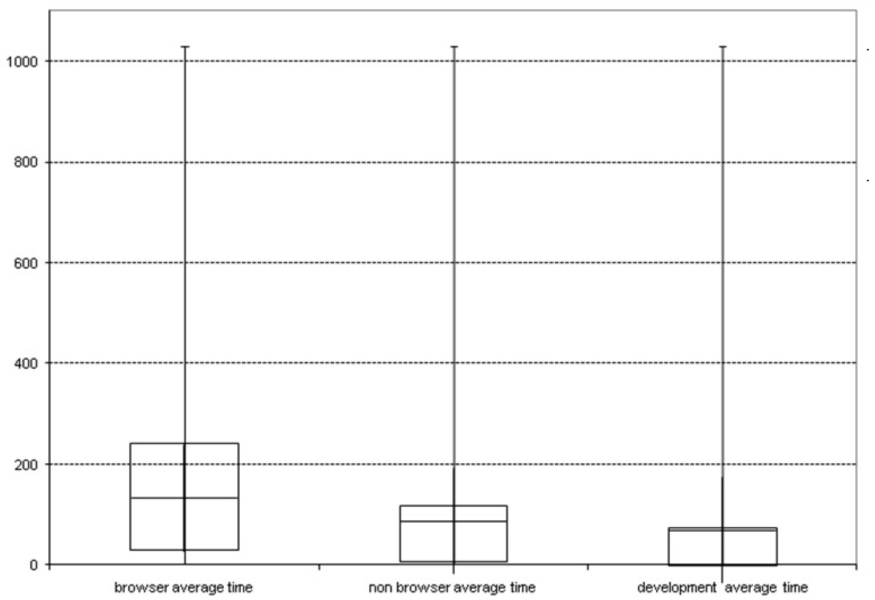


Figure 6.29: Software Developer Active Times - Active times of interaction. Bar on the left represents the sessions during which the information was copied from the browser. The bar on the right represents sessions during which the copy-and-paste was done within the development environment, the middle bar represents the average for all other sessions.

As a consequence of the above Figure 6.29 is a good illustration of the amount of attention that was paid to the development documents depending on the source of the information used. The figure illustrates the extent of differences between those distinct groups of interactions. Further analysis of the numerical data showed the extent of the

	Q1	Q2	Q3	Average
Browser	40.25	133	242.25	203.69
Non-browser	25.25	87	117.25	98.40
Development	24	69	74	74.82

Table 6.11: Levene's test of homogeneity of variance for task carried out in different number of locations

difference in the attention measured through active time.

The analysis of the histograms of the active time distribution clearly indicates their non-normal nature once again suggesting a log-normal distribution. The Kolmogorov-Smirnov test confirmed the assumption of log-normality of the captured data and the Levene's test carried out on the data transformed by natural logarithm supported the hypothesis of homogeneity of variance (F(2,1155)=1.547,p=.213).

Since the assumptions of the one-way ANOVA were met by the data, this test was carried out to investigate the differences between the active times based on the source of the copy-and-pasted information. There was a statistically significant effect of the source of copy-and-pasting (when sourcing information from the browser, development environment and other sources) on the observed active time as determined by the one-way ANOVA (F(2,1155)=48.372,p=.000).

Post-hoc comparisons using the Tukey HSD test (based on the data transformed by a natural logarithm) indicated that the mean score for the active time when the information was sourced from the development environment (= 4.45, = 1.49) and the other source condition (= 4.32, = 1.50). However, the browser and other source condition did not significantly differ from the development environment condition. The differences between the active times of different participants were also analysed using the method discussed above, and since the Levene's test could not support the hypothesis of the homogeneity of variance, it must be concluded that there are differences between the users as far as the distribution of active time is concerned.

Similarly, the users spent more time between the copy-and-paste occurrences when the information was sourced from the browser. Figure 29 illustrates the average frequency of copy-and-pasting across all of the users as well as the magnitudes of the copy-and-paste frequency differences.

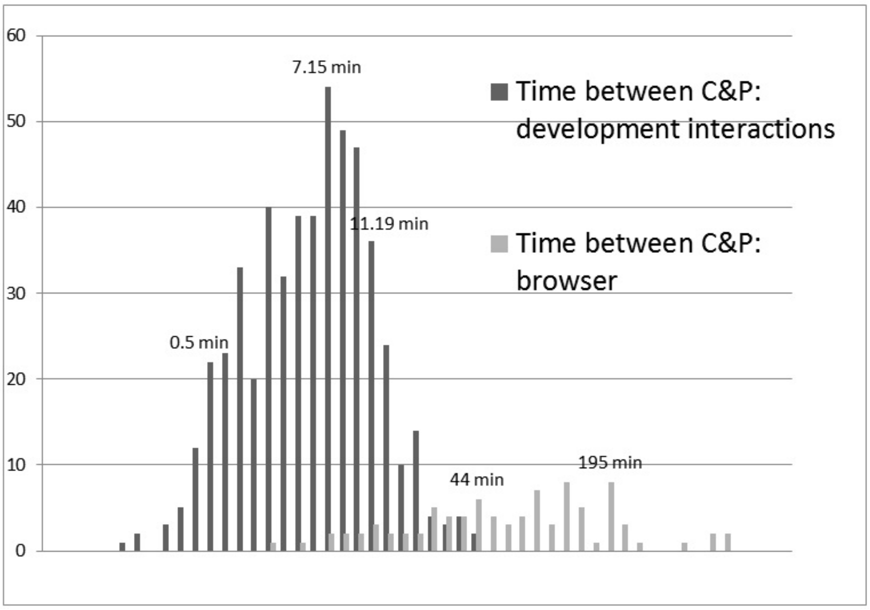


Figure 6.30: Average Time Distribution - The average time between the Copy-and-Paste actions. The darker plot illustrates the distribution of C&P in the development environment. The lighter plot illustrates the distribution of C&P from the browser.

Similarly the analysis revealed the log normal nature of the frequency distribution. The Kolmogorov-Smirnov tests confirmed the assumption of log-normality in all three cases. In contrast to the previously discussed distribution the Levene's test of homogeneity of variance suggested that the three populations had different variances ($F(2,816)=5.548, p=.004$).

In order to assess the behaviour correlation with problem complexity in more detail further analysis was carried out based on the copy-and-pasted content. The data captured during the experiment was processed and each copy-and-paste event was analysed in terms of its content length and subjective complexity measured on a three point scale. The content was encoded with a complexity of zero if it did not carry any logic, a complexity of one if there was logic being copied but is of marginal complexity and a complexity of two if more complex logic was being copied.

Table 6.12 shows the summary of the data analysis. The summary data further supported the observation of higher problem complexity when the interaction through copy-and-pasting from the browser is observed.

The Levene's test was not able to reject the hypothesis regarding the homogene-

Description	Average
Average number of characters copied from the browser	230.05
Average number of characters copied internally in the development environment.	107
Average number of copy and pastes from the browser in a single report	2.24
Average number of copy and pastes from the browser in a single report (the reports having just one copy and paste were excluded from the calculation)	3.86
Average number of copy and pastes internally in the development environment	5
Average logic complexity copied from the browser	0.46
Average logic complexity copied from the development environment	0.20

Table 6.12: Copy-and-pasted content analysis

ity of variance (F(2,1143)=.982,p=.375). The Kruskal-Wallis test showed the existence of a statistically significant difference in complexity between the three distinct populations($\chi^2(2, N = 1146) = 18.782, p = .000$).

Contrasting the complexity of the interaction when the information was sourced from the browser, with the complexity in the group where the information was sourced exclusively from the development environment, revealed a higher complexity of browser based interactions (U(N = 881) = 27177,p ¡ .01). Similarly contrasting the complexity of development environment interactions with all other monitored copy-and-paste interactions revealed higher complexity of other interactions (U(N = 1066) = 92888,p ¡ .005).

The further analysis of the length data and distribution across different sources of copy-and-pasting showed that the three distributions were non-normal and their variance was not homogenous. After transformation with a natural logarithm the hypothesis of the data log normality could be accepted (based on Kolmogorov-Smirnof test). The assumption of data variance homogeneity based on Levene's test could not be accepted either, regardless of the logarithmic data transformation (F(2,1143)=23.466,p=.000).

The above results demonstrate that when copy-and-pasting information from the browser, the users interacted longer with the source code files and much less frequently. Moreover, the data copied from the browser had higher complexity and was longer illustrating a higher complexity of a problem.

The results show that even though copy-and-pasting from the browser was not as frequent as internal development environments operations, it was much more frequent than reported in the feasibility study. The distribution of browser interaction had a normal shape with an average of 120 minutes with Q1=26 min and Q3 = 132 min.

Table 6.12 shows the results of the analysis of the copy-and-pasted text depending on its source. It can be seen from the diagram that the average number of characters that were copy-and-pasted from the browser were more than two times higher than the average number of characters copied from other sources. On the other hand, users copied the information from the browser much less frequently than internally in the development environment. Single copy-and-pastes accounted for 60% of copied data, with the maximum of 8 in a single report.

The data for development environment copy-and-pasting also contained a lot of isolated events with 60% of copy-and-pastes occurring only 1-3 times in the report. The remainder of the reports contained higher values ranging 44-46 indicating a substantial amount of work being carried out on the document.

Finally the relationship between the copy-and-pasting and project building (so the process of code verification, linking and compilation that lead to creation of an executable file) was investigated:

- After the user copy-and-pasted the information from the browser, the build action was executed in 57.5% of cases and with the focus in the same window
- After the copy-and-pasting of the information from the development environment the build took place in 11.14% of cases and with the focus in the same window

In terms of the data source selection during the development activity they were captured through the monitoring of the source process that was used to copy-and-paste the information into the development environment (Figure 6.31).

Figure 6.31 provides an overview of the processes used as a source of the information that was further pasted into the development environment. Due to the policy imposed by the administration and the fact that the IE was the main development target in this experiment it was the most frequently used browser, and a source of copy-and-paste as well as focus-change action (note: program is focused if it currently accepts user input; change of focus means the change of the currently active program). This means that

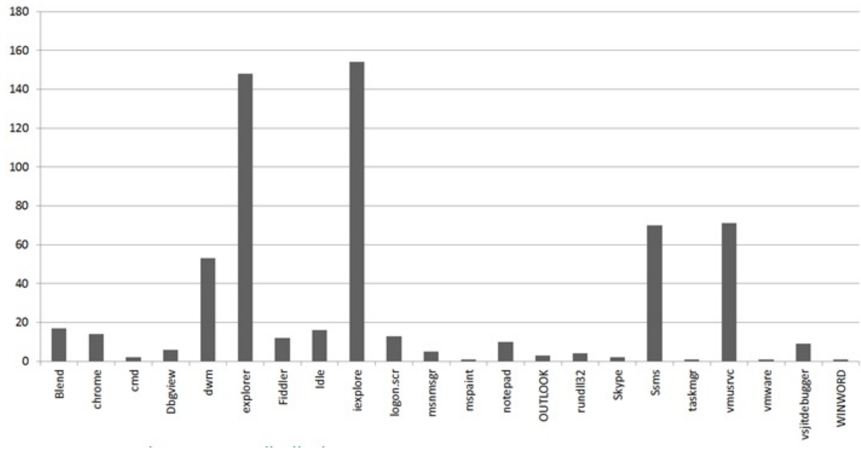


Figure 6.31: Copy-and-paste process distribution -

not only users copied information most frequently from the browser but the interaction with the browser was a primary reason for development interruptions. Figure 6.31 does not show the interaction within the development environment as due to a high count of those interactions (917 non-trivial interactions), the scale of the distribution was obstructed. Tools such as "Blend", "cmd", "Dbgview", "Fiddler", "Ssms" or "vsjitdebugger" are complementary development tools but were not included in the direct observation. The "Vmusrvc" process is a part of Microsoft Virtual PC which was being used extensively during the code development for hosting the development and test environments locally on the machine.

6.4.2.2 Indirect opportunistic practices: Information reuse

This subsection focuses on the second group of shortlisted behaviours reported by the participants of User Study 2. The focus of this analysis is to review the information identification activities during code delivery, which is this case, was measured through the change of the user focus. Since no eye/gaze tracking solution was used in the study due to its ethnographic nature, the information regarding user focus was limited to the information about the active program/document and the interaction of users with the programs/documents. As a consequence the focus-change information was very similar to the data captured while observing copy-and-pasting behaviour wit h the key difference being that the data was transferred between the programs without the assistance of the computer (i.e. directly by the user) and as a consequence, in this analysis the exact content that the user-focused on could not be taken into account.

It was hypothesised that the correlation of user active-time in the environment to the presence and target of focus-changes was similar to the correlation of the copy-and-pasting activity. The analysis of the distribution of the active time once again confirmed its non-normal nature (Kolmogoto-Smirnov p=.000). The Levene's test confirmed the assumption of homogeneity of variance therefore the (F(2,1155)=q.547,p=.213) further non parametric analysis was carried out. The Kruskal-Wallis test showed statistically significant differences between the active times depending on the program group used for focus-changes ($\chi^2(2, N = 1158) = 89.752, p = .000$). The post-hoc paired comparison using Kruskal-Wallis illustrated that the active time when the focus-changes did not occur or occured only internally in the development environment was smaller than when the focus-changes occurred between the browser and development environment ($\chi^2(1, N = 876) = 31.972, p = .000$) or other programs ($\chi^2(1, N = 1068) = 31.972, p = .000$). Even though the mean active time for development interaction was higher for focus-changes taking advantage of the browser, the differences were not statistically significant ($\chi^2(1, N = 372) = .243, p = .622$).

Very similar results can be observed for the focus-change frequency data depending on the focus-change source. Similarly to the previous analysis the data illustrates differences in variances depending on the focus-change source, with the assumption of homogenous variance rejected by the Levene's test (F(2,816)=159.901,p=.000). It has to be concluded that the three groups are indeed different.

6.4.3 Population behavioural consistency

In order to assess the overall behaviour consistency of the group of software developers the experiment was divided into two separate time periods each having the same number of "non-trivial reports" (as defined in section 6.4.2.1). The different observations belonging to the two different time periods were further analysed in order to observe the changes to the distributions and the magnitude of those changes.

The analysis first focused on the complexity variation in time for the development task. The first complexity parameter tested was copied content length. The Kolmogorov-Smirnof test did not support the hypothesis regarding the normal distribution of the length data (p=.000). The Levene's test supported the hypothesis of variance homogeneity between the samples (F(1,1144)=.903,p=.342). The Kruskal-Wallis test did not allow the rejection of the hypothesis regarding the similarity of the two selected populations ($\chi^2(1, N = 1146) = .348, p = .555$) .

Similarly the analysis of the complexity score in time proved not to be normally distributed (Kolmogorov-Smirnof: p=.000) but the variance of the distributions was homogenous (Levene's test:F(1,1144)=.002,p=.967). The Kruskal-Wallis test did not allow the rejection of the hypothesis regarding the similarity of the two selected populations ($\chi^2(1, N = 1146) = .027, p = .871$) .

Secondly the behavioural parameters were investigated. Results similar to the above were obtained when analysing the frequency of the copy-and-pastes distribution of the two time periods studied (Kolmogorov-Smirnof: p=.000, Levene's test:F(1,817)=.074,p=.786, Kruskal-Wallis test: $\chi^2(1, N = 819) = .136, p = .713$) as well as the overall observed active time (Kolmogorov-Smirnof: p=.000, Levene's test: F(1,1156)=.089,p=.765, Kruskal-Wallis test: $\chi^2(1, N = 1158) = 1.997, p = .158$) failing to identify statistically significant differences between the populations.

The distribution of the focus-change count across the two time periods was consistent for all of the users. The data failed the tests of normality but the Levene's test supported the assumption of data variance homogeneity (F(1,1156)=.554,p=.457). The Kruskal-Wallis test failed to reject the assumption of similarity of the focus-change count distribution across the two populations ($\chi^2(1, N = 1158) = .523, p = .470$) .

Similarly the analysis of the focus-change frequency distribution in two intervals failed to identify statistically significant differences (Kolmogorov-Smirnof: p=.000, Levene's test: F(1,817)=.004,p=.950, Kruskal-Wallis test: $\chi^2(1, N = 819) = 1.168, p = .280$) in the two distinct time samples supporting the hypothesis of user and group macro-behaviour consistency.

6.5 Interpretation of the results

This section provides a brief summary of the results discussed in this chapter and provides their interpretation in relation with the research questions.

6.5.1 Extent of HIB consistency

The results clearly illustrated that the distribution of user interaction across geographical locations and time was fairly narrow. The Majority of users interacted with IR systems in a small number of settings and were faced with a limited and recurring number of problems. Even though a big proportion of user tasks could be classified as "one off" or "sporadic" the majority of reported user activity could be classified as highly recurring (i.e. 49% - diary study, 42% - automatic observation and 120 observations per study being the threshold for the tasks classified as recurring) or even habitual (41.9%). Similarly, there was very little variation of user location when interacting with the same tasks with more than 80

The analysis of the remaining moderating-factors revealed that their distribution across different settings, tasks and time, was stable as well. By grouping user responses based on "Context" (as defined in this study) and Task Classification the Coefficient of Variation of all of the moderating-variables dropped by an order of magnitude (e.g.: stress level 72.58% - 7.10%, search query formulation difficulty 62.71% - 15.33%).

The above results confirmed the proposed hypothesis on the stability of user interaction context and task (discussed in Chapter 4). The analysis suggested that users tend to interact in the same locations, during the same periods of time, with the same tasks.

Similarly the behavioural responses to recurring tasks were consistent. Both the self-reported (e.g. throughout the User Study 1 and 2), as well as the observed (e.g. in User Study 3) behavioural distribution functions, had a non-uniform nature supporting the hypothesis of consistency/rationality in human information choices. Clear trends of user behaviour could be identified throughout all of the measured parameters.

This does not mean that the nature of HIB consistency was trivial. On the contrary, a variance of user behaviour was apparent. Variance occurred across different user tasks, geographical locations, emotional states, goals, perceived difficulty levels and a wide range of factors captured during User Study 1. A substantial variability was observed also between the users. This made the generalisation of behavioural results difficult. This supported the conclusions of the discussion regarding the limitations of moderating-factor approaches.

On the other hand the data analysis illustrated that certain properties of HIB could be used to improve the prediction accuracy. The results illustrated that the data variance could be considerably decreased if user interactions were grouped based on the user/group, location, time of interaction, and task type. This supported not only the assumption of consistency as outlined in this thesis but also supported the validity of the proposed framework.

The results also suggested that during IS&R humans develop information-related habits and that some proportion of user behaviour could be supported by automated cognitive processes. The results reported in this study are consistent with the more general cognitive research and suggested that the cognitive processes involved are the same.

To summarise, the results supported the hypothesis of user environmental and behavioural consistency.

6.5.2 Identification of consistent behaviours

The results clearly illustrated that there existed a finite set of highly repeatable and consistent behaviours which were triggered as a response to successful information retrieval. Moreover the above mentioned behaviours were executed frequently enough to be considered a systematic source of implicit or semi-implicit relevance-feedback. The above results indicate that there existed a group of widely reoccurring interactions correlated to information relevance:

- Interaction with code examples in parallel with code development.
- Copy-and-paste of code snippets at a block, subsystem or system level.
- Source code, library or system download attempts.
- Any of the above followed by code experimentation.

All of the identified behaviours relate to user opportunistic/development by example practices.

The analysis carried out in User Study 2 suggested that the following are the key factors which affected the attitude towards the identified behaviours and social norms, as well the perception of control:

- The stage of the project: The results illustrated that the stage of the project moderated the user role in the project, as well as the user preference towards the information objects and their retrieval frequency. The impact of the stage of the project on the observed behaviour was not direct. It was further moderated through the task/role. Multiple users reported that their interaction frequencies varied drastically depending on the stage of the project.

- Task / Role/ Archetype: The results supported the observations made previously by Janjic et al. (2010) or LaToza et al. (2006) and clearly illustrated that both macro as well as micro-behaviour is moderated by the task/behavioural archetype. It is important to note that multiple users in this study reported that their roles varied considerably across the entire spectrum of development tasks. Participants were engaged in scoping, development, test and post-release support activities and their behaviour varied as these activities changed. As one of the users stated: interaction frequency was "Highly dependent on the stage of the project. At the early stages the interaction was very intensive and most of the work Internet based. In the latter stages of the project the level of interaction drops dramatically, sometimes to the extent when the Internet was not needed at all".

- Size and complexity of the information object: The self-reported frequencies of user interaction illustrated that the reuse of small elements of code (code snippets) were reported to have the highest frequency. The size of code also affected the perception of control and also indirectly the social acceptance of the opportunistic processes (as the concern over licensing and quality issues related differently with elements of code of different size). This was observed both quantitatively and qualitatively through the behavioural frequencies captured. The qualitative analysis was based on the interpretation of user comments (e.g. the fact that they engaged in opportunistic practices "mainly for very small problems where the solution is clearly defined").

- Time and Location of HIB. Participants reported that the above activities were strictly work-related although some reported frequent interaction with customers at customer sites and sometimes at home as well. The reported variation related mainly to the location of interaction. Although the majority of interactions were reported to take place during the standard working hours (where high variations of working hours were recorded), where there was a need, some of the work was performed outside of working hours.

The results support the hypothesis of user behavioural-response consistency by reinforcing the observation from the literature review and exposing clear and rational development-related information strategies.

6.5.3 Identification of moderating-factors

The results illustrated that the identified HIB was affected by a number of moderating-factors:

- Work/Non-work-related nature of the task/context: This was observed through a bimodal distribution of the captured data. The analysis revealed that the coefficient of variation of data decreased when the data was split into two separate populations based on this moderating-factor.
- Time and location: This factor had an impact on user behaviour and selection of tasks. The relationship between the task, location and time was very strong and stable for all of the participants. The distribution functions did not vary substantially throughout the time of experiment.
- Individual differences: Significant variations were observed between the participants. The variations did not affect the interpretation of the identified behaviours but affected their overall frequency and distribution in time and also across the tasks.
- Stage of the project: The analysis of population consistency did not show statistically significant variations of user behaviour in time.
- Task / Role / Archetype of behaviour: The analysis of the retrieved content did not allow for identification of any variations of behaviour caused by the user role or tasks type. The type of the task impacted the way the development environment was being used. The task affects the decision whether to use the environment or not and the amount of the actual development/copy-and-paste behaviour.

- Size/Complexity / Commercial nature of the problem: -

The above were the only factors significantly affecting the shortlisted HIB. The analysis failed to reveal statically significant change to the observed behaviours throughout the duration of the experiment. Moreover the analysis failed to identify statistically significant variations of the identified moderating-factors as well.

6.6 Conclusions

The results analysis presented in this chapter supports the research hypothesis of HIB consistency. The analysis supports the findings of this thesis and illustrates that even though HIB varied considerably across numerous moderating-factors the users in the studies tended to interact with the same tasks in the same context/location and over the same periods of time. From the perspective of the Black Box Model this preference of the users to use stable environments minimised the moderating-factor variation. This as a consequence minimised the HIB variation as well and allowed the research to reduce the complexity of the research problem.

The outlined "Integrated Framework for HIB in-situ" was successfully applied to the problem of work tasks of software developers and led to the identification of behaviours that could be used in IR development. It also allowed for the identification of the key moderating-factors affecting the variability of user HIB and for understanding the impact the variability has on the IR system.

The next chapter concludes this thesis, presents its strengths and weaknesses and discusses the further research that will be carried out.

Chapter 6: Summary

User Study 1 revealed that a big proportion of user activities was highly consistent/repeatable (49%). The consistency can also be observed across all of the input variables of the General Black Box model.
User Study 1 also revealed that 41.9% of user activities met the definition of habitual behaviours by Wood, Quinn, and Kashy. The observed habitual behaviour was consistent with the general habitual behaviour supporting the hypothesis of the consistency of cognitive mechanisms responsible for IS&R and everyday activities.
The behavioural analysis which was carried out through the User Study 2 revealed high potential of the shortlisted behaviours to predict relevance. The results illustrated very high and positive attitude towards the identified behaviours (participants perceived the behaviour as a natural consequence of finding relevant information). Similarly the level of perceived behavioural control, especially for smaller information objects was reported to be very high. In a majority of cases the identified behaviours were highly accepted in the work group.
Further analysis of shortlisted behaviours (User Study 3) of software developers identified only the following factors as having statistically significant impact of the behaviour: • Work/Non-work-related nature of task along with the task time and location • Individual differences between the users • Stage of the project • Task / Role / Archetype of behaviour • Size / Complexity / Commercial nature of the problem
User Study 3 did not identify any statistically significant changes to the observed behaviours or the above mentioned moderating-factors throughout the period of the study further supporting the hypothesis of human behavioural consistency.
Successful application of the proposed framework supports the validity of the "Integrated Framework for HIB in-situ".

Table 6.13: Summary of chapter 6

7

Conclusions

7.1 Contributions to knowledge

The work described in this thesis makes a number of contributions to the existing knowledge by investigating the long acknowledged problem of consistency in HIB through user studies. Importantly this thesis has proposed a systematic framework for contextualised/personalised information retrieval system. The framework is based on two well-known theories: Theory of Planned Behaviour and Holistic Cognitive Theory of Information Retrieval. Notably, the effectiveness of the framework is demonstrated by applying this to a case study of software engineers. The rest of the chapter discusses the main contribution in detail and also the perceived limitations and strengths are highlighted in the chapter.

7.1.1 Consistency of Human Information Behaviour

The key contribution of this thesis is the investigation of the HIB consistency. There exists a significant research in HIB and IS&R investigating human behavioural variability and factors affecting it. However this research is lacking in understanding of cognitive mechanisms of human consistency. On the other hand psychological research focuses on the problem of human behavioural consistency but is not easily applicable to the HIB and IS&R problem. The work presented in this thesis bridges those two research domains. It illustrates how user everyday cognitive mechanisms take part in IR and it builds foundation for HIB consistency.

The results of the work presented in this thesis clearly support the hypothesis of HIB consistency. This does not mean that the human consistency is a simple problem. The results of this work indicate clearly that in order to effectively predict behaviour it is not sufficient to obtain knowledge about a limited number of moderating-factors but to analyse the context of the problem, user and interaction.

Results of User Study 1 reveal that:

- Users tend to make similar decisions about the similar sets of tasks as long as they are in very similar contexts.
- Users intentionally choose to interact with the same sets of tasks.
- Users intentionally choose to interact in the same contexts.
- User behaviour changes significantly across tasks.
- User behaviour changes significantly across contexts.
- User behaviour changes significantly with the individual differences between the users.
- For every user there exist a small number of highly repeatable tasks which are being carried out in highly repeatable contexts and as a consequence lead to highly repeatable behavioural responses.
- The above mentioned tasks and contexts do not change significantly in time.
- User response to the above mentioned tasks does not change significantly in time.
- It is possible to identify those tasks, contexts and behaviours using the proposed "Integrated Framework for HIB in-situ"

7.1.2 Integrated Framework for HIB in-situ

The second key contribution of this thesis is the integrated framework which is developed based on two well-known approaches: Theory of Planned Behaviour and Holistic Cognitive Theory of IR. The framework helps to bridge the gap between the cognitive research which offers effective tools for obtaining high predictive validity when investigating human behaviour and the IS&R research which focuses significantly on the investigation of HIB and development of IR utilities. The proposed framework helps to systematically design personalised and contextualised IR solutions as discussed in detail in chapter 4.

7.1.3 Consistency of software developers' Information Behaviour

The third key contribution to knowledge is the application of the proposed framework to the problem of the work-related information retrieval of software engineers and the consistency of their information behaviour. The application of the proposed framework to the group of software engineers further confirmed the initial hypothesis of HIB consistency and the applicability of the framework to IS&R problems.

The application of the framework to the work-related problems of software engineers resulted in:

- Identification of the actor group responsible for the highest IR related cost: Software Developers and Testers
- Identification of the tasks responsible for the highest IR related cost: Development, Debugging and Maintenance of Code.
- Identification of the highly reoccurring behaviours directly correlated to information relevance:
  - Direct code reuse practices at block level (distribution of cognition)
  - Direct code reuse at subsystem and system level (opportunistic programming)
  - Direct reuse of system level components e.g. libraries, classes, frameworks and bigger components (component reuse)
  - Indirect code and information reuse practices at different complexity levels (block, subsystem, system)
- Assessment of the level of consistency of the identified behaviours and identification of the factors responsible for the majority of variance through user attitude analysis.

The micro behaviours proved to be not only very good implicit indicators of relevance but also frequent enough to allow for their widespread use in IR personalisation and contextualisation.

7.1.4 Identification of strong implicit-feedback indicators for the population of software developers

The attitude analysis carried out in User Study 2 allowed for the identification of moderating-factors affecting the variance of the shortlisted behaviours. The impact of the identified moderating-factors was then verified in User Study 3. User Study 2 used the proposed framework to investigate:

- The attitude towards opportunistic practices: The results illustrate that the overall attitude towards the copy and pasting and knowledge reuse is very favourable. Developers perceive Internet as a very valuable source of code related information and use it as the first information resource.
- The subjective norm: Opportunistic practices overall are perceived as a valuable method for accelerating the development process. As long as they do not impact the overall quality of the developed solution, or do not have legal implications are perceive as socially acceptable or even encouraged.
- Perception of control: Developer feels comfortable using on-line resources. They admit that sometimes finding the solution to their problem is difficult and often proves to be a challenge but they feel in full control of the search process. Apart from the social and economic restrictions there are very little factors affecting the controllability of the search process. As a consequence the level of perceived control is highly correlated to the size and complexity of the work problem.

The analysis illustrated that because of high levels of attitudes towards opportunistic practices, lacks of negative social norms and high perception of control, opportunistic behaviours are a frequent response to relevant information especially for smaller problems where those levels are even higher. The analysis carried out in User Study 2 revealed a number of moderating-factors that are affecting the extent to which the developers engage in opportunistic behaviours, including:

- Work/Non-work-related nature of the task/context
- Time and location
- Individual differences
- Stage of the project
- Task / Role / Archetype of behaviour
- Size/Complexity of the problem

The above were found to be the only factors to have a statistically significant effect on the opportunistic behaviours.

7.2 Strengths

One of the strengths of the research discussed in this thesis is its very strong theoretical as well as empirical foundation. The review of cognitive and IR literature carried out in this work clearly illustrates the consistency of the cognitive mechanisms involved in the everyday as well as information-related problem-solving and decision-making. The thesis, through the user studies, provides additional empirical evidence supporting author's claim that the mechanisms responsible for human use of information are the same as those responsible for all other decision-making and problem-solving tasks. The experimental evidence supports this claim and gives high confidence that the results of this work are indeed generalizable to other domains, user tasks and contexts. The ability to support the results of this work with such a vast body of research give this work significant theoretical advantage over other models of HIB and research design methodologies.

Secondly, the use of the proposed framework, allows systematic investigation of the level of behavioural variance explained by the identified moderating-factors. Obviously the amount of variance that can be explained is limited and will indeed vary across different tasks and contexts but the historical research as well as the results of this work demonstrate that it is possible to predict more than 20%-30% of variance without resorting to complex analysis and the introduction of any moderating-variables.

Thirdly the work presented in this thesis provides a very clear and simple example of how the proposed framework can be applied to the selection of the population, the set of tasks, as well as the context in which the interaction takes place. Moreover the analysis of software developer behaviour leads to very clear and concise conclusions regarding the task-specific system design and implementation.

The thesis proposes a comprehensive framework entailing its steps and processes involved in its application. Importantly the effectiveness of this framework is demonstrated by applying this to a case study of software developers.

7.3 Weakness

The perceived limitation of the framework lies in the fact that it focuses on very narrow user groups, tasks and contexts. Such a narrow focus of the framework simplifies the design process considerably and allows for drastic reduction of behavioural variance and number of moderating-variables to be taken into consideration. On the other hand the results of the application of the framework from one domain, task and context are not easily transferable to a different domain, task or context. One cannot simply take the results of the work analysis of software developers performing their coding activities and transfer it to (for example) a context of human resources workers. Those two user groups work in a significantly different way and will be engaging in different activities and different information behaviours will be repeatable/consistent. The results of framework application cannot even be easily transferable to other tasks that a wider group of software engineers is carrying out such as technical specification design, or preparing for workshops with customers. Even though both of those tasks might involve a number of consistent behaviours most probably they will not be the same. The research process (application of the proposed framework) for those activities has to be repeated. Provided there is sufficient financial viability the process of redesign and redevelopment of the end IR application has to be carried out again. The proposed framework is highly generalisable though, as it is based on very basic and universal cognitive mechanisms involved in the human decision-making process.

Another perceived limitation is related to the sample used for the behavioural analysis of developers' behaviour. The sample consisted of very experienced software developers highly specialised in the use of .NET technologies and using only a single family of development tools. The results were not verified on other software developer groups or on groups with significantly lower and/or higher expertise. This is a direct consequence of time limitation of this thesis and the entire project as well as very high cost of removing the human resources from their normal activities for the purpose of experimentation. This is a part of the future work described in the next section.

Finally, the proposed framework cannot be treated as a development process. Transformation of this framework into an easy and intuitive development process requires a significant amount of work including documentation, preparation of training materials and utilities. This work is significantly beyond the scope of this thesis.

7.4 Future work

The work carried out in this thesis is multidisciplinary and not only covers many areas of research from cognitive psychology and IS&R research, but also owing to the source of the funding, was significantly business driven. The work carried out during this study revealed a significant gap not only between the Library Sciences and IR, which is well acknowledged (Ingwersen and Järvelin, 2005*a*; Saracevic et al., 1997), but especially in the wider context between Cognitive Psychology and IR. Modern cognitive theories describing the problem-solving and decision-making process of human beings are very well established and there exists a very significant potential for their application in IR. This refers not only to the application of existing cognitive theories to the UI and workflow design but also to the implementation of personalised IR and IS&R solutions.

The future work following up this research can be classified depending on the timescale of the planned effort:

- Current Work: For almost a year now the author of this research has been employed by Jaguar Land Rover research. At the moment the proposed framework is applied or partially applied to 3 research projects focusing on information management and user behavioural profiling. One of which is Technology Strategy Board funded and carried out in collaboration with Coventry University. The second project is expected to be supported by further EPSRC funding (subject to EPSRC approval).

  At the moment of writing there is ongoing work for a patent publication and there is already a paper awaiting publication for the patent submission.

- Short Term:

  The work which was carried out for the purpose of this thesis and Case Studentship resulted in a significant amount of cross-disciplinary literature that was reviewed and organised. Similarly the description of the application of the methodology to the group of software developers as well as the findings of current projects (once protected by patents), are being collated and are planned to be disseminated.

  There is also a discussion taking place aiming to elect a number of PhD students to follow the proposed methodology in their research in order to investigate other IS&R domains and problem spaces. The above include the development of personalised-systems for software development students and potential implementation of implicit-feedback based enterprise contact-exchange systems.

- Long Term:

  The long term plan as well as the author's personal ambition is to transform the proposed framework into an easily applicable design methodology. And this ambition is close to becoming a reality as a significant number of software components that can be used together to deliver search and information management utilities are already freely available on an open source or commercial basis. Using solutions such as Apache Solr/Lucene, Apache ManifoldCF, Apache Tika, Apache Hadoop and Mahout search utilities can be very quickly and effectively developed. What is missing is a methodology that would hide from the developers the complexity of the personalisation and contextualisation problem, and help them focus on the key development activities.

  In order to achieve the above there still needs a significant amount of effort to be invested. In order to allow software development companies to develop effective solutions as discussed in this thesis, without the need for a costly research process, there is a need for the creation of an entire ecosystem of materials and support, including:

  - Further publications discussing empirical results of research containing: experimental case studies, processes employed during research, research findings, proposed system implementations and analysis of real-life system performance.

– Creation of further tools, software utilities and documents supporting the process and allowing non-research staff to gather basic information about user groups and their activities as well as information-related behaviours.
– Creation of supporting training documentation (books/manuals) and organisation of community focused support resources (website, blogs, Wikipedia entries) as well as commercial support and consulting services.
– Dissemination not only in the IS&R community but also in the significantly wider software-development world. What is even more important is the dissemination across companies who are struggling with the complexity of information management. There is a need to systematically inform key business stakeholders about the size and nature of the information management and retrieval problem, as well as inform them about a wide number of solutions which can be applied to the problem.

Hopefully, carrying out the above activities will allow for the systematic development of highly-personalised IS&R utilities by non-research specialised software engineers and significantly simplify the design and development process.

Chapter 7: Summary

The results of User Studies supported the hypothesis of HIB consistency. Furthermore they have supported the hypothesis of consistency of software engineers' information behaviour.
Human consistency is highly dependent on a significant number of moderating-factors and high levels of predictive validity can currently only be achieved through the analysis of clearly defined actors carrying out repeatable tasks in fixed contexts.
The proposed "Integrated Framework for HIB in-situ" proved to be effective in identifying the consistent behaviours as well as identifying the moderating-factors responsible for a big proportion of behavioural variance.
The proposed framework was successfully applied to the problem of HIB of software engineers. It allowed for an identification of a number of behaviours highly correlated to relevance as well as the factors moderating this behaviour.
The proposed framework is very task/actor/context sensitive (hence the name: HIB in-situ). The framework is highly generalisable and because its foundations in "Theory of Planned Behaviour" it can be applied to any task/actor/context group. The results of the application of the framework to a certain group are specific to that group only and can not be transferred to a different group of without further reapplication of the framework.
Future work focuses on the further implementation of software developer focused information-management utilities and the application of the proposed framework to other information-demanding actor groups/ tasks.

Table 7.1: Summary of chapter 7

Appendix A

User Study 1 Materials

A.1 Introduction

This appendix presents the key research materials used during the User Study 1. The appendix provides copies of the:

- Screening questionnaire
- Checklist used during the interview
- The diary material provided to the experiment users
- Checklist used for the purpose of the closing interview
- Miscellaneous documents including the experiment-consent and ethics-related documentation

A.2 Screening questionnaire

The goal of the screening questionnaire was to ensure that the study participants elected for the study meet very limited study requirements (mainly ensuring that the participants actually use the Internet). The secondary goal of the questionnaire was to capture basic demographic information about the users in order to get a better picture of the population.

Screening Questionaire

The goal of this questionnaire is to:

- get some basic information about you and
- verify whether you fit the required profile for this research

1. Name?
2. Surname?
3. Email?
4. Phone number?
5. Age?
6. Gender (please tick)? Male ☐ Female ☐
7. What is your present occupation?
8. Give more details about your current job (if student give more details about the curse you are taking)?
9. Does your job involve using a computer (tick if true)? ☐
10. Do you have an internet connection at work (tick if true)? ☐
11. Do you have an internet connection at home (tick if true)? ☐
12. How many hours daily do you use internet?
    a. I do not use it on a daily basis (specify)?
    b. Less than 1? ☐
    c. Between 1 and 2? ☐
    d. Between 2 and 5? ☐
    e. More than 5? ☐

Figure A.1: Screening Questionnaire 1 - Page 1 of the User Study 1 screening questionnaire

13. Out of all the time that you spend using your computer what percentage do you spend doing the following?
    a. Creating or reading documents (word, excel, power point, pdf, etc...)?
    b. Computer games?
    c. Internet for professional reasons (university, work)?
    d. Internet for private reasons?
    e. Other professional?
    f. Other private?
14. Select a sentence that describes you best?
    a. I am a proficient computer user. Both the hardware as software architecture is known to me to some extent. ☐
    b. I am an advanced computer user. I do not know the details of computer architecture but I can manage my computer on my own. I know how to use the operating system and most of the software available. ☐
    c. I am an average computer user. I know how to use some basic software like a word processor or a spread sheet but I leave administrative task to specialists. ☐
    d. I am a new computer user. ☐
15. Have you ever been frustrated about the quality of the search results you receive? Is yes select how often?
    a. Seldom or never? ☐
    b. Occasionally? ☐
    c. Frequently? ☐
16. Have you ever had problems with formulating a search query that would precisely describe your information needs? If yes select how often?
    a. Seldom or never? ☐
    b. Occasionally? ☐
    c. Frequently? ☐
17. If you do not find the satisfying results within the first several results returned by the search engine do you (select as many answers as you wish)?
    a. Redefine the search query? ☐

Figure A.2: Screening Questionnaire 2 - Page 2 of the User Study 1 screening questionnaire

b. Look on the next pages of the search results? ☐

c. Go to first several pages and look for the related links? ☐

d. Search using another search engine? ☐

e. Give up? ☐

18. Which search engine do you use(select as many answers as you wish)?

a. Google ☐

b. Live Search (MSN) ☐

c. Yahoo ☐

d. Other (specify)?

19. Why did you choose the search engine selected above and not any other?

Figure A.3: Screening Questionnaire 3 - Page 3 of the User Study 1 screening questionnaire

A.3 Interview Checklist

The goal of the check-list was to guide the interview process and to provide rich materials that were designed to aid the interaction between the interviewer and the interviewee. When discussing the interview process it is important to note the three goals that the interview was intended to achieve:

- The primary goal of the interview was to capture initial and basic understanding of the way people interact with information resources. The checklist was used during the interview to try to analyse the key information routines of the user. A calendar captured the date and time where interactions occurred, which was supported by a map showing the particular location. Through the process of the interview and all the information captured from the checklist, calendar and map the rationale for the choice of location and the reasons for the choice could be established. This review would effectively allow the behaviour and the factors affecting the behaviour to be analysed.

- The secondary goal of the interview was to ensure that participants understood the questions and there were no doubts around the experimental process. In order to make sure that the experiment documentation was clear it was first reviewed with a group of 3 independent participants.

- The third goal of the interview was to ensure that the diary information was captured in a wider range of information-interaction contexts. During the interview the first page of the diary was populated (interaction map) and the key places where internet access occurred were clearly marked. Study participants were also encouraged to provide diary entries also to record the locations.

1. Interview Checklist

Participant ID... Interview Place..

Interview Date.. Interview Time...

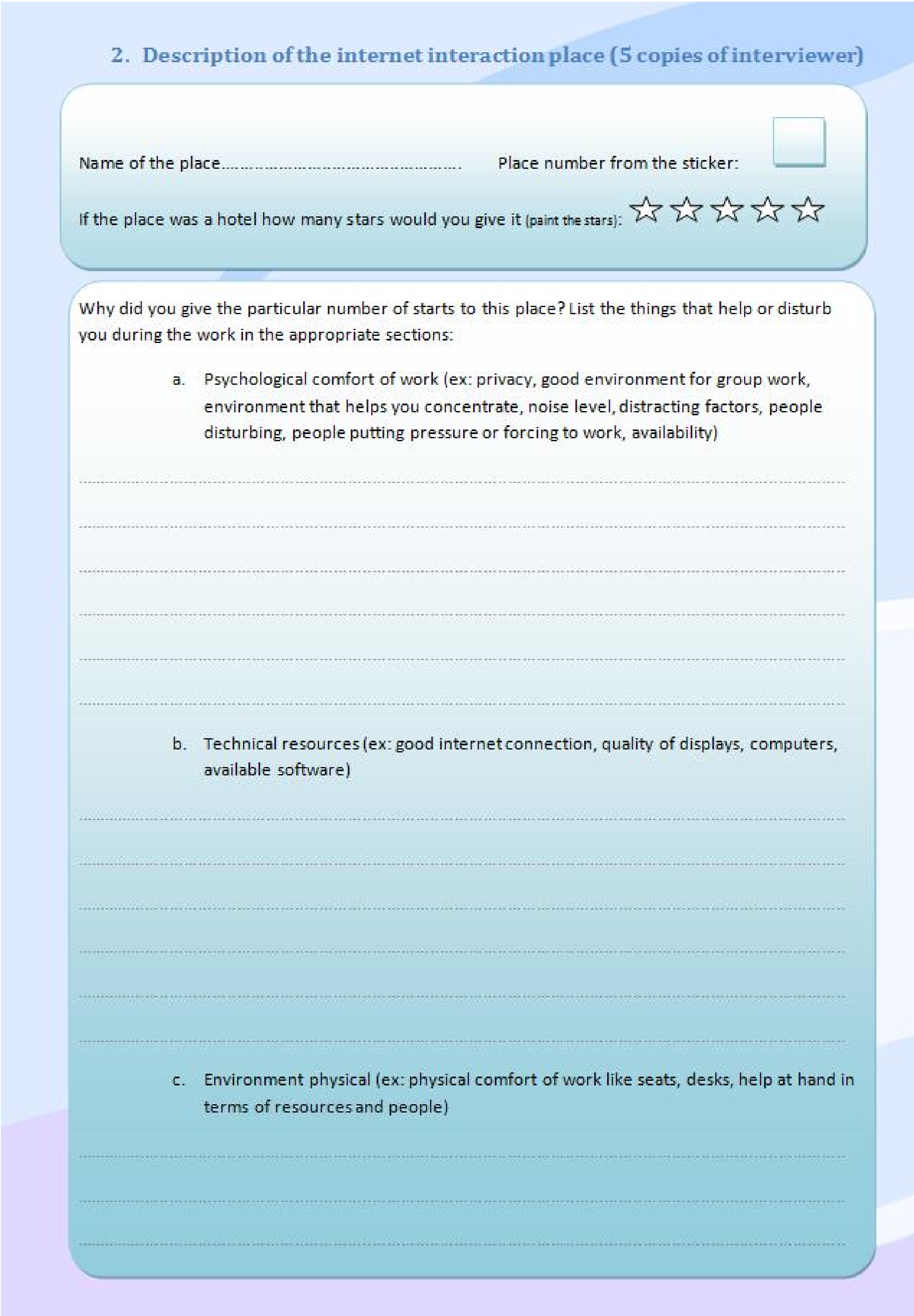
1. Details of participation explained ☐
2. Need for recording explained ☐
3. Recording started ☐
4. Payment details explained ☐
5. Award for the most active participant explained ☐
6. Contract explained and signed ☐
7. User code specified on the contract ☐
8. Online account created ☐
9. Passwords specified on the contract and the research diary ☐
10. Consent form explained and signed ☐
11. Diary explained and handed in ☐
12. Interview about the place done ☐
13. Debriefing meeting discussed ☐
14. Example of entries discussed ☐
15. Exemplary search done and documented ☐
16. Details of debriefing written on the diary and this document ☐

Debriefing Interview date: ....

Debriefing interview time: ....

Debriefing interview location: ....

Figure A.4: Interview Checklist 1 - Page 1 of the User Study 1 interview check-list



2. Description of the internet interaction place (5 copies of interviewer)

Name of the place... Place number from the sticker:

If the place was a hotel how many stars would you give it (paint the stars): ☆ ☆ ☆ ☆ ☆

Why did you give the particular number of starts to this place? List the things that help or disturb you during the work in the appropriate sections:

a. Psychological comfort of work (ex: privacy, good environment for group work, environment that helps you concentrate, noise level, distracting factors, people disturbing, people putting pressure or forcing to work, availability)

b. Technical resources (ex: good internet connection, quality of displays, computers, available software)

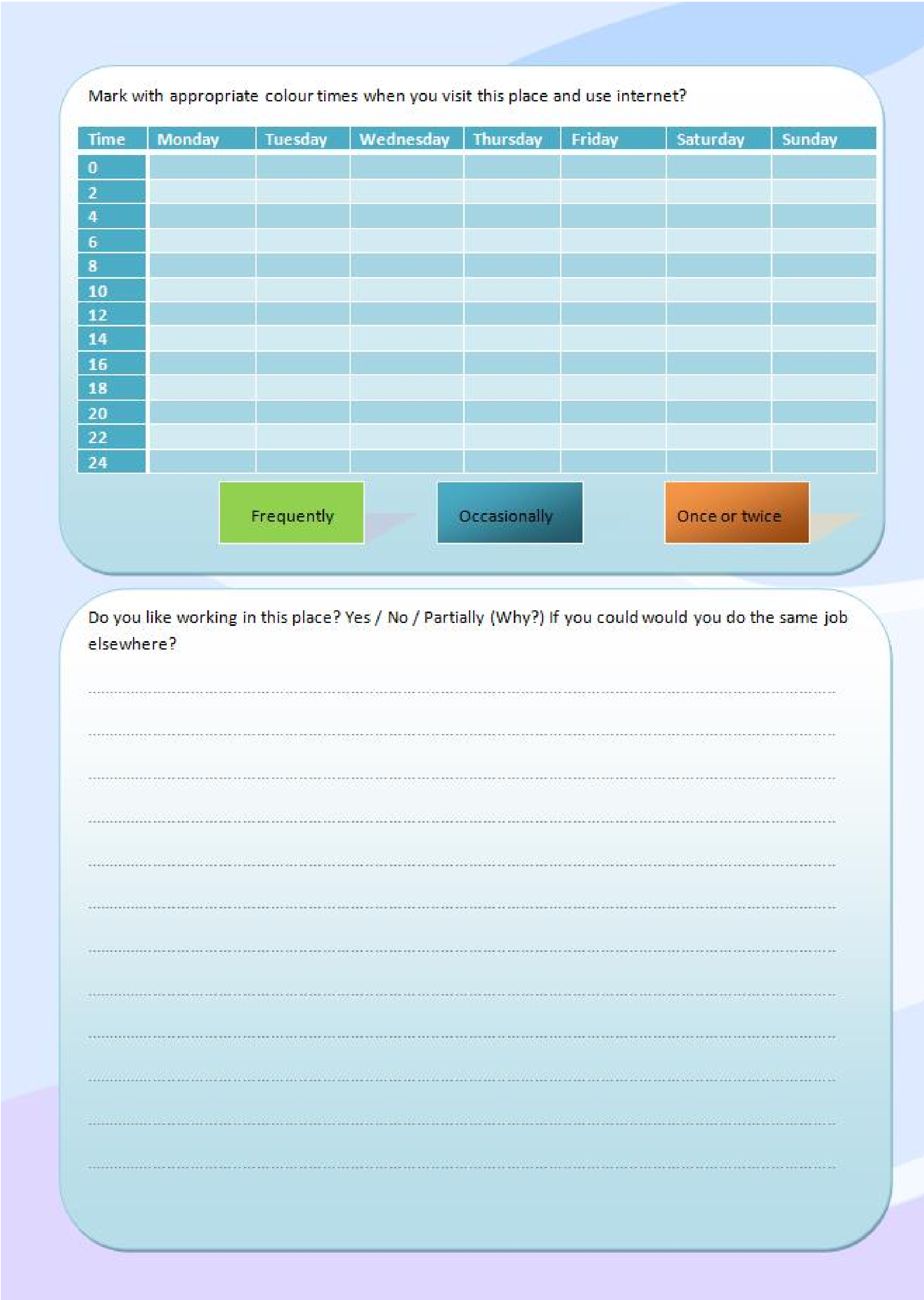
c. Environment physical (ex: physical comfort of work like seats, desks, help at hand in terms of resources and people)

Figure A.5: Interview Checklist 2 - Page 2 of the User Study 1 interview check-list

d. Other

List the things that you most frequently do on the internet in this place? (Do you search for this particular information also in other places and how frequently? Why do you search for it in this particular place?)

Figure A.6: Interview Checklist 3 - Page 3 of the User Study 1 interview check-list



Mark with appropriate colour times when you visit this place and use internet?

Time	Monday	Tuesday	Wednesday	Thursday	Friday	Saturday	Sunday
0							
2							
4							
6							
8							
10							
12							
14							
16							
18							
20							
22							
24							

Frequently Occasionally Once or twice

Do you like working in this place? Yes / No / Partially (Why?) If you could would you do the same job elsewhere?

Figure A.7: Interview Checklist 4 - Page 4 of the User Study 1 interview check-list

A.4 Diary Study

The diary was completed by the participants in their free time during the everyday or work-related information interaction. Some of the sections though were filled-in during the interview. The goal of those sections, apart from maintaining the user information was also to discipline the action of the user to ensure that the tasks performed as well as the location where it occurred were properly recorded.

Since a significant number of internet activities were highly repeatable (e.g. checking on-line email clients, reading the on-line news or writing on Facebook) those activities were captured through single diary entries. Participants were asked to only annotate the repeatable entries with dates/times of interaction.

Similarly a proportion of the activities spanned a significant amount of time and accounted for a very high number of search queries. For those activities participants were encouraged to write on the back side of the diary page with the detailed description of the activity intentionally left blank.

It is also important to note that the extract below does not show the full diary content which for the purpose of the study contained multiple copies of pages 6 to 9.

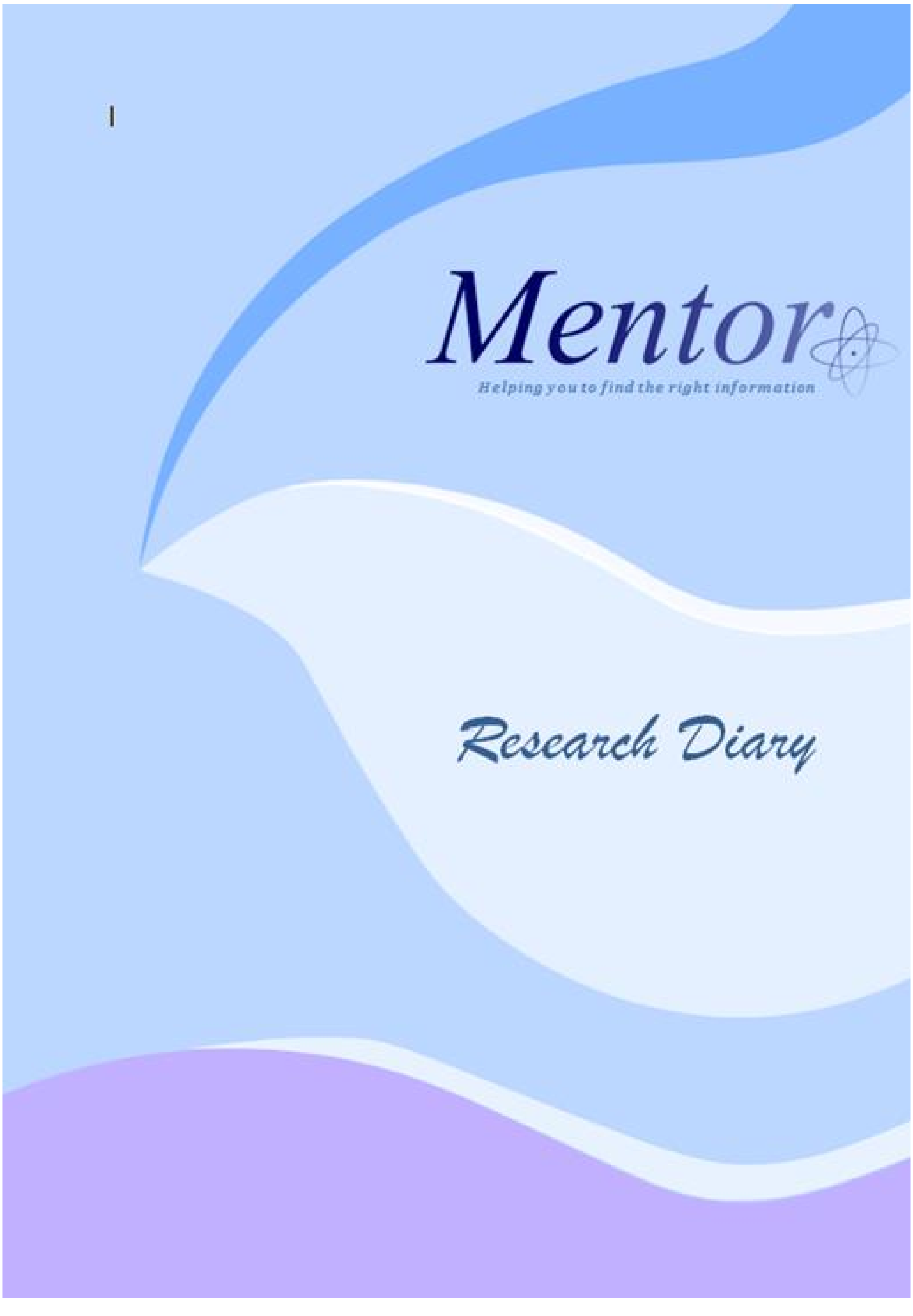


Figure A.8: Diary Study 1 - Page 1 of the User Study 1 diary materials

Thank you for taking part in this user study,

The goal of this research is to investigate how people look for information on the Internet. The aim of this user study is to observe how people choose the search strategies and to what extent they are consistent in choosing them.

During this experiment you will be asked to describe the information you are looking for and describe the ways you are looking for it.

We hope that thanks to your participation we will be able to develop new improved search interfaces and techniques.

If you have any doubts or questions feel free to contact:

Adam Grzywaczewski	aa4747@coventry.ac.uk	AS233
Rahat Iqbal	R.Iqbal@coventry.ac.uk	AS218

Beast Regards

The research team

Figure A.9: Diary Study 2 - Page 2 of the User Study 1 diary materials

Participant ID: ....

Project address: http://mentor.coventry.ac.uk/Experiment1

Date of the interview: ....

Place of the interview: ....

Debriefing Interview date: ....

Debriefing interview time: ....

Debriefing interview location: ....

Figure A.10: Diary Study 3 - Page 3 of the User Study 1 diary materials

A.5 Closing Interview

The goal of the closing interview was to ensure that the entire process had been performed correctly and to gather additional feedback and user observations. Since the users were compensated for participation in the studies, there was a reward for the most comprehensive diary and the financial/ contractual matters also had to be completed during the final interview. This involved checking the quality of the diary and briefly reviewing the content and any process-related difficulties.

1. Your Internet Places

Using the stickers provided mark the places where you usually use internet. Use appropriate sticker for appropriate device. If the location is not on the map place the sticker next to the description field below. If you have any doubts have a look at the detailed instructions for section 1.

Campus:

Unknown places:

1 Name of the place: ....

2 Name of the place: ....

3 Name of the place: ....

Figure A.11: Diary Study 4 - Page 4 of the User Study 1 diary materials

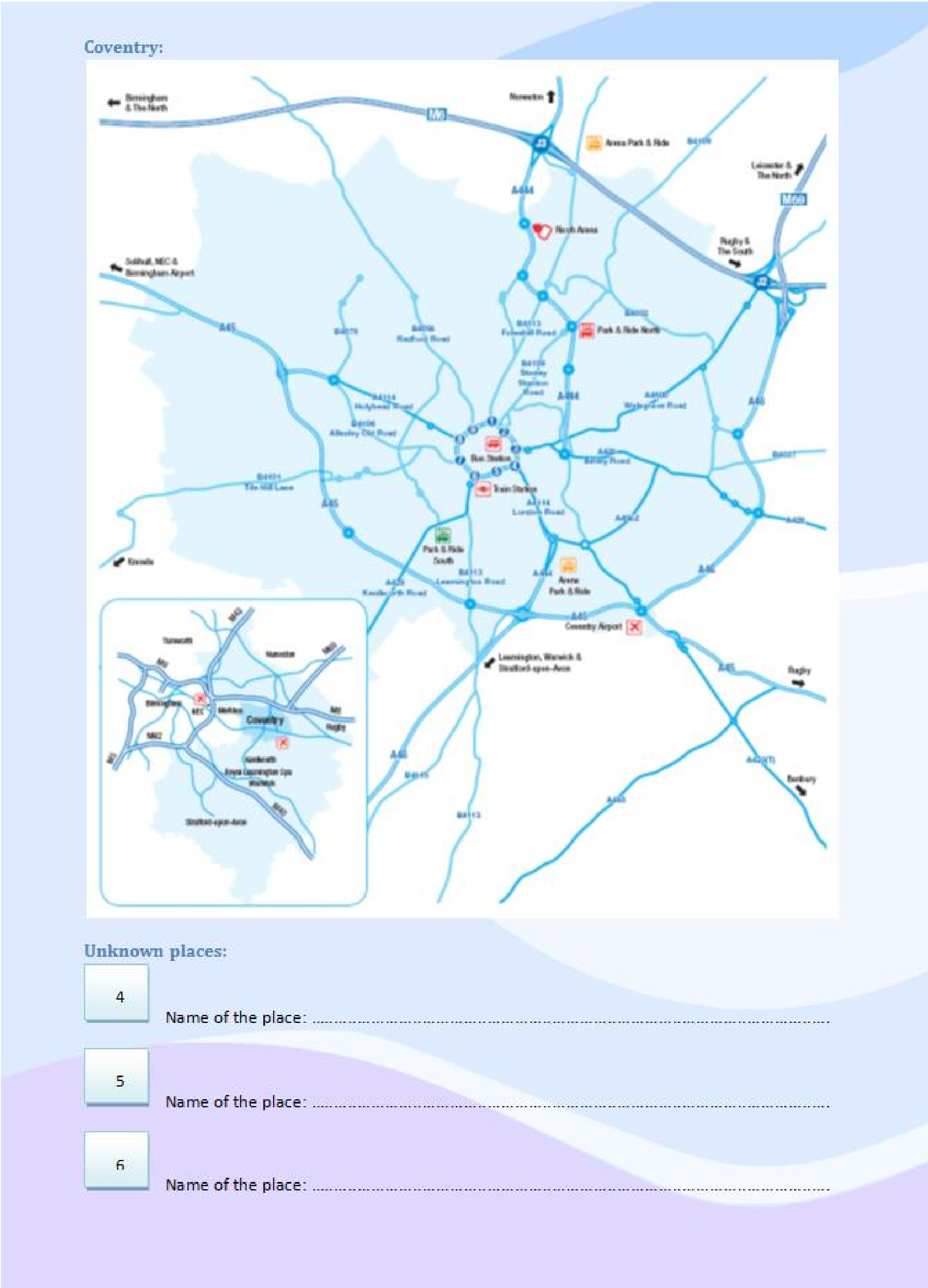


Figure A.12: Diary Study 5 - Page 5 of the User Study 1 diary materials

2. Internet Activities Diary

Whenever you start a new search on the Internet answer the following questions describing your internet interaction.

Day number........... Entry number......... Number From the Sticker ....

Date: ....

Where are you working? : ....

Time: ....

Working alone or in group? :................................

Write briefly what information you have been searching for.

....

....

....

....

How would you classify the information?

- The information is work/university related. ☐
- You look for information for pleasure. ☐
- Every day personal duties (paying bills, buying groceries). ☐
- Other (if you feel that the ones above don't match the information you are looking for)

....

....

Why do you need the information for?

....

....

....

....

....

....

Figure A.13: Diary Study 6 - Page 6 of the User Study 1 diary materials

Select the description of the search task that best matches the task that you have been performing (you can mark more than one if you wish):

- "Fact Finding Task" (looking for predictable information for example: book in the library, course mark, buss schedule) ☐
- "Information Gathering Tasks" (looking for unknown information for example: looking for information for the research, looking for a job, looking for new wireless card) ☐
- "Browsing Tasks" (going through known or unknown information without a clearly defined goal for example: going through the favourite blog, going through news, playing videos) ☐
- "Transaction Tasks" (tasks in which you are performing an online action for example: web based email, messaging, interaction with the bank website) ☐
- Other kind of task (if you feel that the task you are performing does not match the above write briefly what is it and how you would categorise it).

Describe the task that you have been performing using the categories listed below.

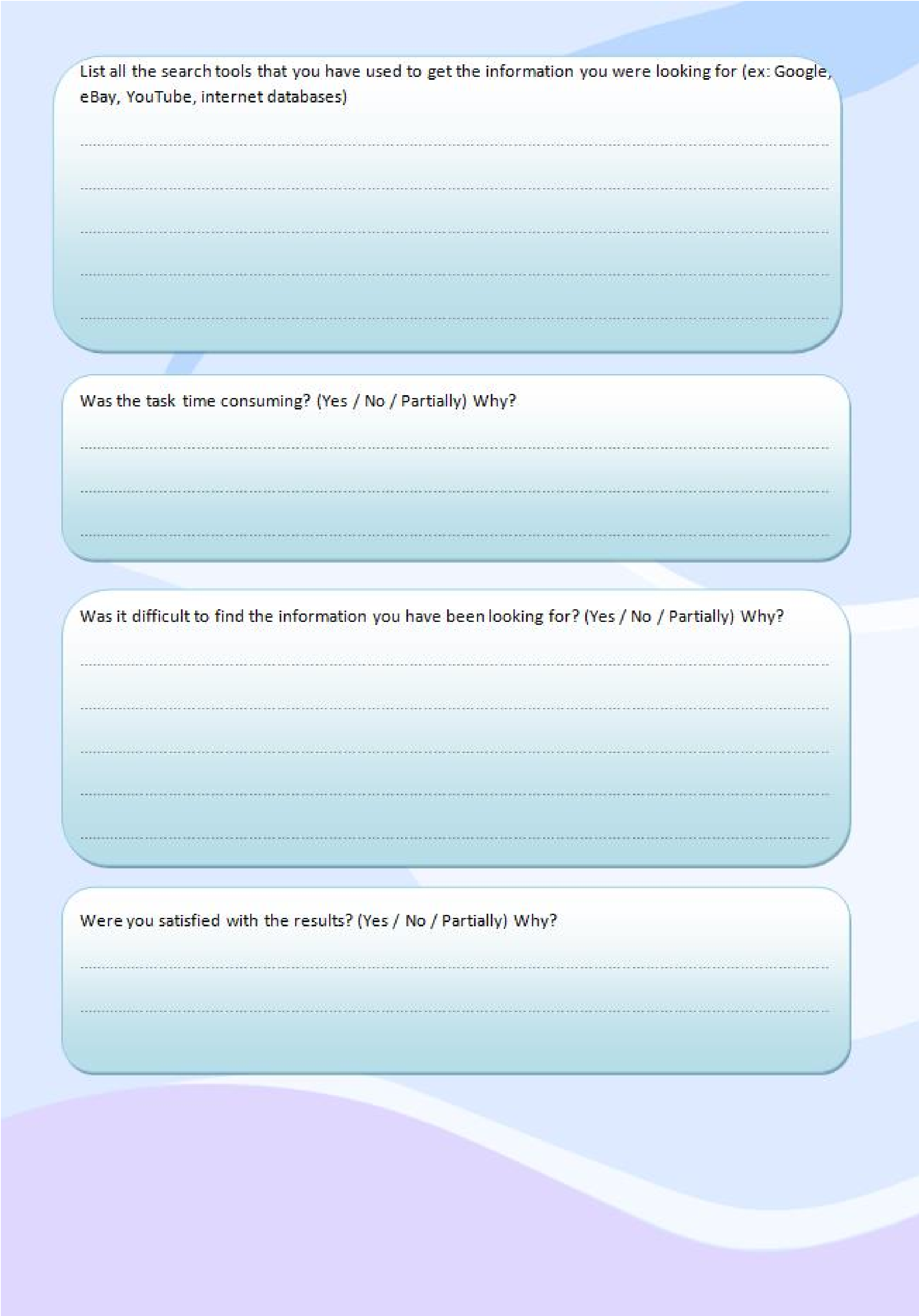
1. Task difficulty (give marks from 1 to 5 where 1 is trivial and 5 is very difficult):
2. How much effort do you think the task will require from you (1 I will find it immediately after I enter the search query or address to 5 it will require multiple search repetitions and a lot of reading):
3. How difficult is it for you to define the search query (1 I know exactly what to write in the search engine 5 it is very difficult for me to define it):
4. Your domain knowledge of the task (1 if you know nothing about the task and 5 if you are a domain expert):
5. Degree of urgency of task (1 no time pressure to 5 highly importance):
6. Your confidence level (1 I don't know whether I will be capable to finish to 5 it is just a matter of time):
7. How important is the information for me (1 the task has no impact on me 5 it is crucial to get the correct information):
8. Stress level (1 no stress to 5 highly stressful):
9. How frequently do you perform this task (write more or less how many times per week):
10. Since when you have been performing the task (write more or less when you have started):

Figure A.14: Diary Study 7 - Page 7 of the User Study 1 diary materials

Write step by step how did you look for this information?

List all the search queries (keywords) that you have used to search for information in the correct order?

Figure A.15: Diary Study 8 - Page 8 of the User Study 1 diary materials



List all the search tools that you have used to get the information you were looking for (ex: Google, eBay, YouTube, internet databases)

Was the task time consuming? (Yes / No / Partially) Why?

Was it difficult to find the information you have been looking for? (Yes / No / Partially) Why?

Were you satisfied with the results? (Yes / No / Partially) Why?

Figure A.16: Diary Study 9 - Page 9 of the User Study 1 diary materials

1. Interview Checklist

Participant ID……………………………………… Interview Place……………………………………

Interview Date……………………………………… Interview Time……………………………………

1. Diary collected ☐
2. Diary verified ☐
3. Account Number Obtained ☐
4. USB Stick given ☐
5. Questions asked ☐
6. Transfer confirmed ☐
7. Personal details deleted ☐
8. Tell about next stage ☐
9. Award verified ☐
10. Award paid ☐

Figure A.17: Closing Interview 1 - Page 1 of the User Study 1 closing interview

2. Closing Interview

What were your impressions about the experiment?

Did you find any of the questions confusing or difficult?

The word task was frequently used in this research, how do you understand it?

Did you perform any task in more than one place?

Figure A.18: Closing Interview 2 - Page 2 of the User Study 1 closing interview

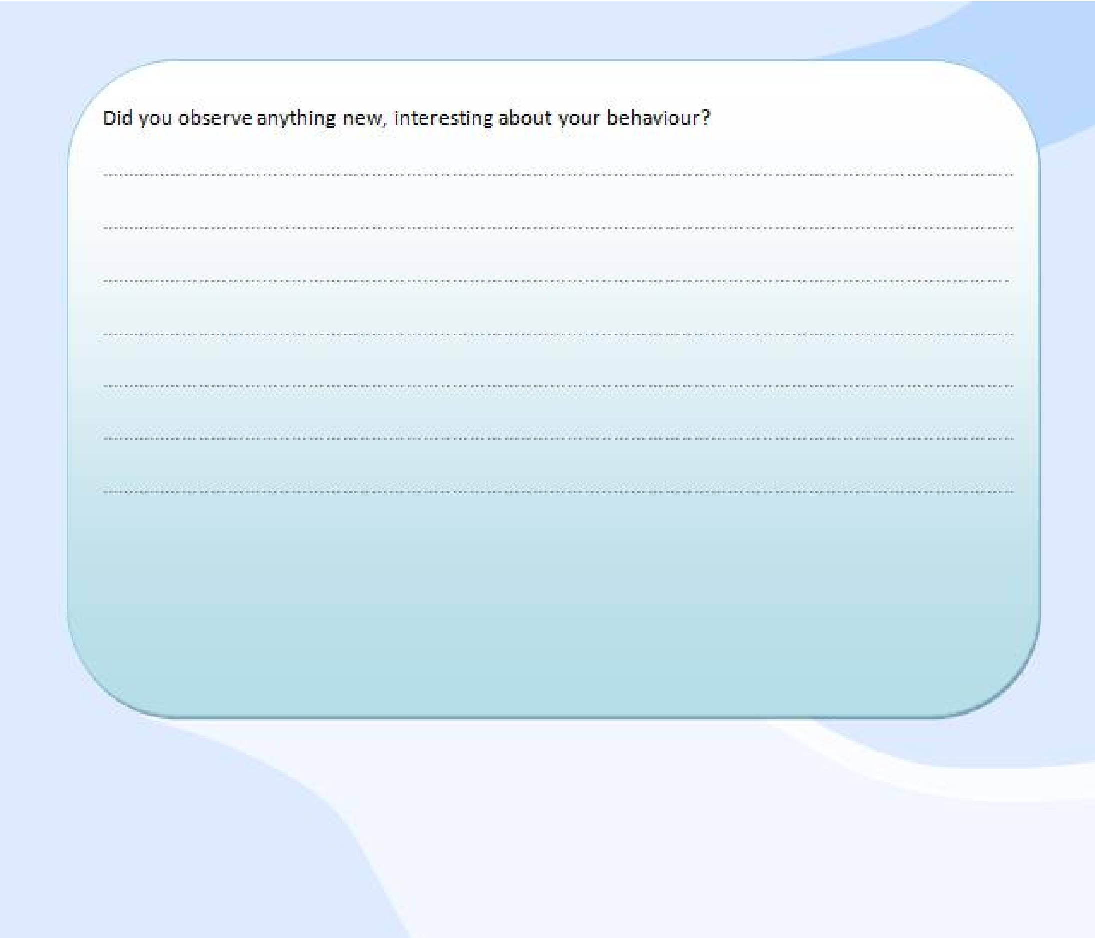
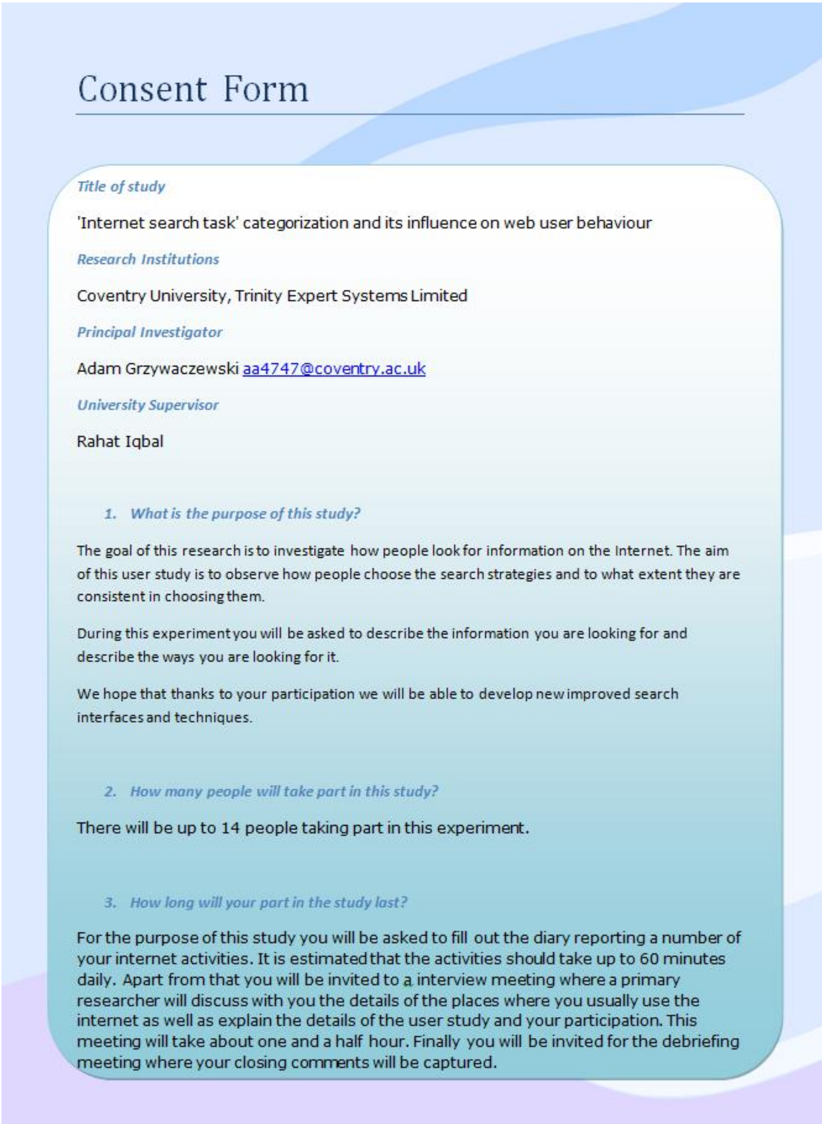


Figure A.19: Closing Interview 3 - Page 3 of the User Study 1 closing interview

A.6 Consent form

The attached consent form was a mandatory part of the Coventry University ethical approval process.



Consent Form

Title of study

'Internet search task' categorization and its influence on web user behaviour

Research Institutions

Coventry University, Trinity Expert Systems Limited

Principal Investigator

Adam Grzywaczewski aa4747@coventry.ac.uk

University Supervisor

Rahat Iqbal

1. *What is the purpose of this study?*

The goal of this research is to investigate how people look for information on the Internet. The aim of this user study is to observe how people choose the search strategies and to what extent they are consistent in choosing them.

During this experiment you will be asked to describe the information you are looking for and describe the ways you are looking for it.

We hope that thanks to your participation we will be able to develop new improved search interfaces and techniques.

2. *How many people will take part in this study?*

There will be up to 14 people taking part in this experiment.

3. *How long will your part in the study last?*

For the purpose of this study you will be asked to fill out the diary reporting a number of your internet activities. It is estimated that the activities should take up to 60 minutes daily. Apart from that you will be invited to a interview meeting where a primary researcher will discuss with you the details of the places where you usually use the internet as well as explain the details of the user study and your participation. This meeting will take about one and a half hour. Finally you will be invited for the debriefing meeting where your closing comments will be captured.

Figure A.20: Consent Form 1 - Page 1 of the User Study 1 consent form

4. What will happen if you take part in this study?

During the initial meeting you will be asked to provide the researcher with your bank account details and sign the required documents such as this consent form, agreement for participation and user coding form protecting your identity. The details will be used by the University to process the payment for the participation. All data gathered about you will be processed using the coding assigned during the first meeting. During this meeting you will be introduced to the details of the experiment.

During the course of the study you will be asked to fill out a diary describing your normal internet usage as well as the environment in which you usually interact. You will have software installed on your machine gathering measurements of your behaviour such as mouse click frequencies and number of visits on sites. You will be asked to submit the diary to the research group and afterwards come for an hour meeting where you will discuss your findings and perform some exercise. The meeting will take up to an hour. You will also be asked to submit the data captured by the software and give back all equipment given to you for the purpose of the study. You can chose to leave at any point of the meeting. The meeting will be video recorded for reference and to help in analysing the findings of the meeting.

5. What are the potential benefits from being in this study?

Hopefully the results of this study will help to identify a generic and topic independent way of search task classification. On the basis of this research another research will take place that will look into the potential dependencies between human behaviour and a category of search task that he is performing. Hopefully this research will help to understand the potential of personalized and context aware search and in the long term improve the existing search algorithms.

6. What are the possible risks or discomforts involved from being in this study?

The experiment does not introduce any risks to the participants.

7. How will your privacy be protected?

The identity of the participants will be gathered for the registration and payment issues only. The identity of participants will not be disclosed neither used in the study. You will be assigned a code number and identified using the code number throughout the entire research. If there will be a need to publish voice or video recordings of the discussion, both the image and sound will be processed to protect your privacy.

8. Will you receive anything for being in this study?

You will be paid according to the conditions of the participation agreement.

9. Will it cost you anything to be in this study?

There will be no costs involved.

10. What if you have questions about this study?

Figure A.21: Consent Form 2 - Page 2 of the User Study 1 consent form

You can contact the leading researcher listed on the top of the document or his academic supervisor.

11. What if you have questions about your rights as a research participant?

In case of any queries or complains you can contact the Ethics Committee of Coventry University.

Participant's Agreement:

I have read the information provided above. I have asked all the questions I have at this time. I voluntarily agree to participate in this research study.

____ _________________

Signature of Research Participant Date

____

Printed Name of Research Participant

____ _________________

Signature of Person Obtaining Consent Date

____

Printed Name of Person Obtaining Consent

Figure A.22: Consent Form 3 - Page 3 of the User Study 1 consent form

Contract of participation

1. Title of study

'Internet search task' categorization and its influence on web user behaviour

2. Research Institution

Coventry University, Priory Street, Coventry, CV1 5FB, United Kingdom

3. Participant of the research

As specified in the section called "***Participant's contact details***"

4. Principal Investigator

Adam Grzywaczewski aa4747@coventry.ac.uk room: AS233

5. University Supervisor

Rahat Iqbal R.Iqbal@coventry.ac.uk room: AS218

6. The purpose of this contract

The goal of this contract is to specify the scope of the experiment that the participant will take part in as well as the details of the payment that will be made to the participant and the criteria of payment.

The contract is signed between the primary investigator and the participant of the research.

7. The purpose of the study

The goal of this research is to investigate how people look for information on the Internet. The aim of this user study is to observe how people choose the search strategies and to what extent they are consistent in choosing them.

During this experiment you will be asked to describe the information you are looking for and describe the ways you are looking for it.

We hope that thanks to your participation we will be able to develop new improved search interfaces and techniques.

8. Payment for participation

All participants will be given £30 and an 8 GB USB memory stick for the active participation in the user study.

Additionally the most active participant of the research will be rewarded an amount of £50.

Figure A.23: Consent Form 4 - Page 4 of the User Study 1 consent form

In a case where participant's contribution in terms of number of entries or quality of entries is very low and the results cannot be used in the research the principal researcher reserves the right to deny the payment to the participant.

The payment will be done to the participant via a bank transfer to the specified account not later than one month after the debriefing session.

9. *Description of the user study*

The user study will consist of the following parts:

9.1. Introductory meeting

The introductory meeting will last about one and a half hour. During the introductory meeting the following actions will be taken.

- Participant will be presented with details of participation listed in this contract.
- The participant will be asked to sign the contract accepting the terms of the contract as well as the consent form.
- Participant will be assigned a unique code and identified using this particular code throughout the user study.
- Participant will be interviewed about the places on internet interaction and their details. The interview will be recorded and the recordings will be encoded and used for further analysis.
- Participants will be demonstrated the exemplary diary entry and instructed how to maintain the diary.

9.2. Diary study

The diary study will last for 14 days. During the study the participants will be asked to fill the entries in the web or paper diary describing their everyday internet activities following the instructions given during the first meeting.

9.3. Debriefing meeting

During the debriefing session participants will handle in the research diary as well as take part in the closing interview summarising their observations regarding the research. This session will last about an hour.

10. *Participant's privacy*

The identity of the participants will be gathered for the registration and payment purposes only. The identity of participants will not be disclosed neither used in the study. You will be assigned a code number and identified using the code number throughout the entire research. If there will be a need to publish voice or video recordings of the discussion, both the image and sound will be processed to protect your privacy.

Figure A.24: Consent Form 5 - Page 5 of the User Study 1 consent form

11. Participant's details

Name: ______________________________

Surname: ______________________________

Address: ______________________________

____

____

____

Date of birth: ______________________________

Telephone number: ______________________________

Email: ______________________________

Participant's Unique Code: ______________________________

Participant's Web Password: ______________________________

12. Participant's agreement

I have read the information provided above. I have asked all the questions I have at this time. I voluntarily agree to participate in this research study.

______________________________ ______________

Signature of Research Participant Date

____

Printed Name of Research Participant

______________________________ ______________

Signature of Person Obtaining the Agreement Date

____

Signature of Person Obtaining the Agreement

Figure A.25: Consent Form 6 - Page 6 of the User Study 1 consent form

Appendix B

User Study 2 Materials

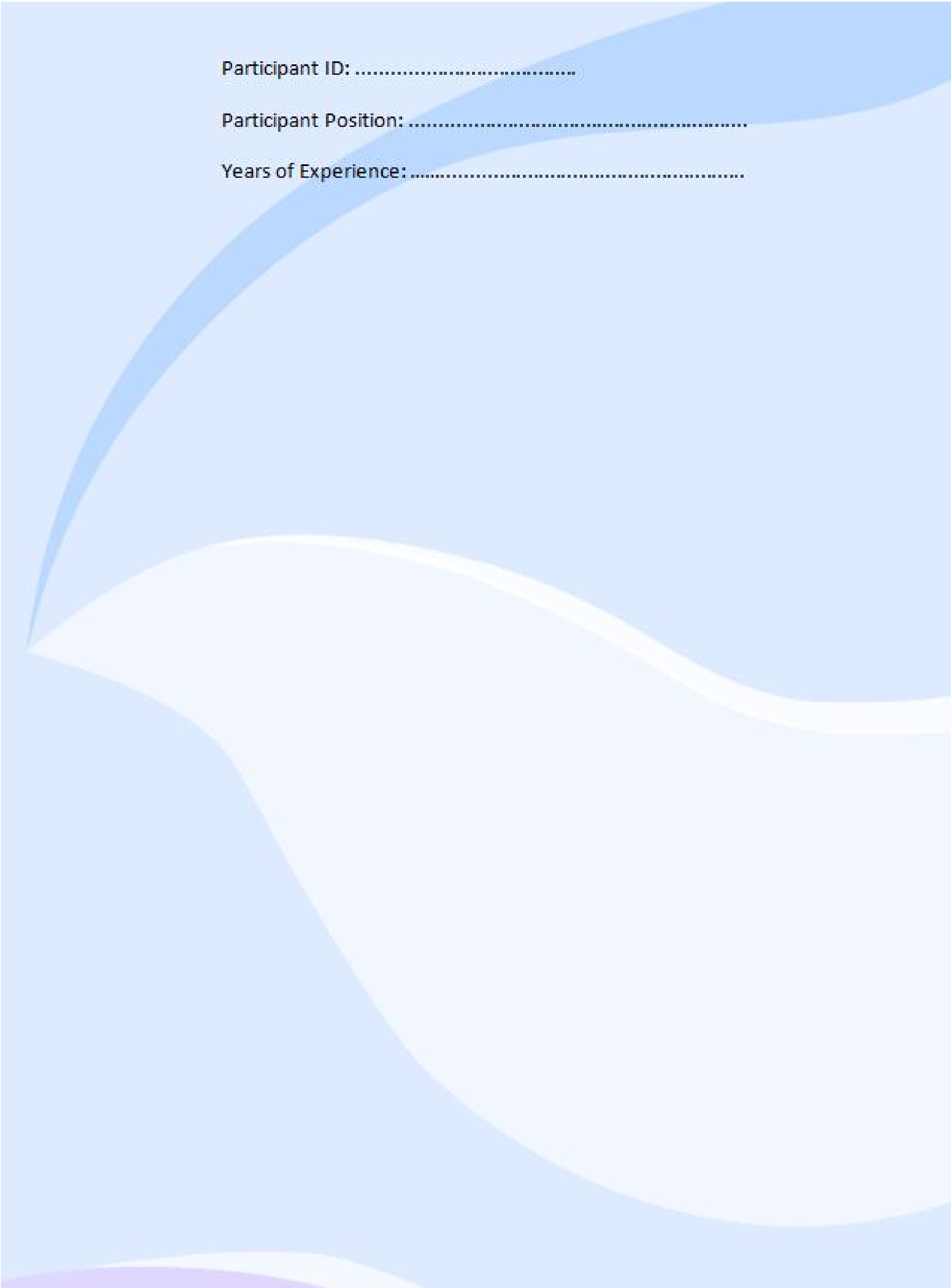
B.1 Introduction

This appendix provides experiment materials used during the User Study 2. It provides the copies of:

- Experimental questionnaires
- Materials supporting the interview process
- Experimental consent documentation

B.2 Questionnaire

The questionnaire in contrast to the screening questionnaire in the User Study 1 played a primary role in User Study 2. Even though it captured the key attributes of the experiment participants the goal of the questionnaire was to reveal and confirm the key initial assumptions about software developer information behaviour extrapolated from the literature review. The questionnaire was supposed to be a quick and inexpensive method for confirming key HIB and create solid foundations for further work, in this case the User Study 2 interview.



Participant ID: ....

Participant Position: ....

Years of Experience: ....

Figure B.1: Questionnaire 1 - Page 1 of the User Study 2 questionnaire

Questionnaire

1. How often your work related problems require from you the use of Internet\Intranet? (select one)
   a. Every week
   b. Every day
   c. (1-5) times in a day
   d. (1-10) times in a day
   e. More than 10 times a day
   f. Other

2. What work related information do you look for on the internet? (select the ones that match)
   a. Code documentation.
   b. Code examples.
   c. Solutions to some code related problems
   d. General ideas regarding the approach to take
   e. Technological news
   f. Training material
   g. Other

Figure B.2: Questionnaire 2 - Page 2 of the User Study 2 questionnaire

3. How often do you look for code documentation and code examples? (select one)
   a. Every week
   b. Every day
   c. (1-5) times in a day
   d. (1-10) times in a day
   e. More than 10 times a day
   f. Other

4. Once you find the information you need what do you usually do with it? (select the ones that match, try to think about last two or three times you have been looking for the information)
   a. Just read it and use the knowledge
   b. Copy the code snippet or code example
   c. Read the information in the same time solving the problem
   d. Print the information
   e. Bookmark, email or save the information in any way
   f. Share the information with other people
   g. Other

Figure B.3: Questionnaire 3 - Page 3 of the User Study 2 questionnaire

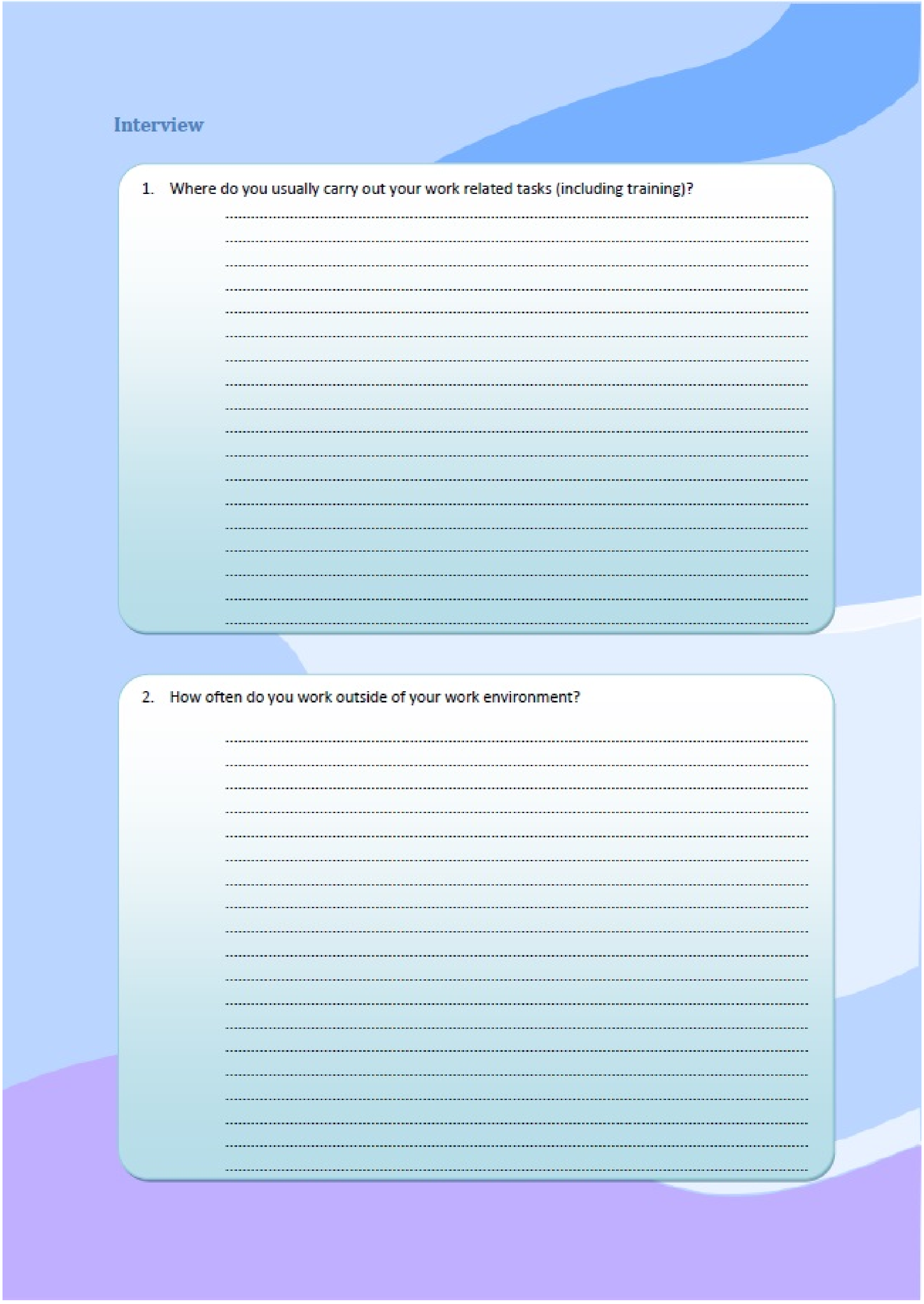
5. How often do you reuse the code snippets of code from external sources to rework them in your projects?
    a. Every week
    b. Every day
    c. (1-5) times in a day
    d. (1-10) times in a day
    e. More than 10 times a day
    f. Other

Figure B.4: Questionnaire 4 - Page 4 of the User Study 2 questionnaire

B.3 Interview

The goal of the interview was to revisit the key methodological questions and provide additional insight into the information retrieval and management process to guide further IR tool development process.

Figure B.5: Interview Schedule 1 - Page 1 of the User Study 2 interview schedule


Interview

1. Where do you usually carry out your work related tasks (including training)?

2. How often do you work outside of your work environment?

Figure B.6: Interview Schedule 2 - Page 2 of the User Study 2 interview schedule

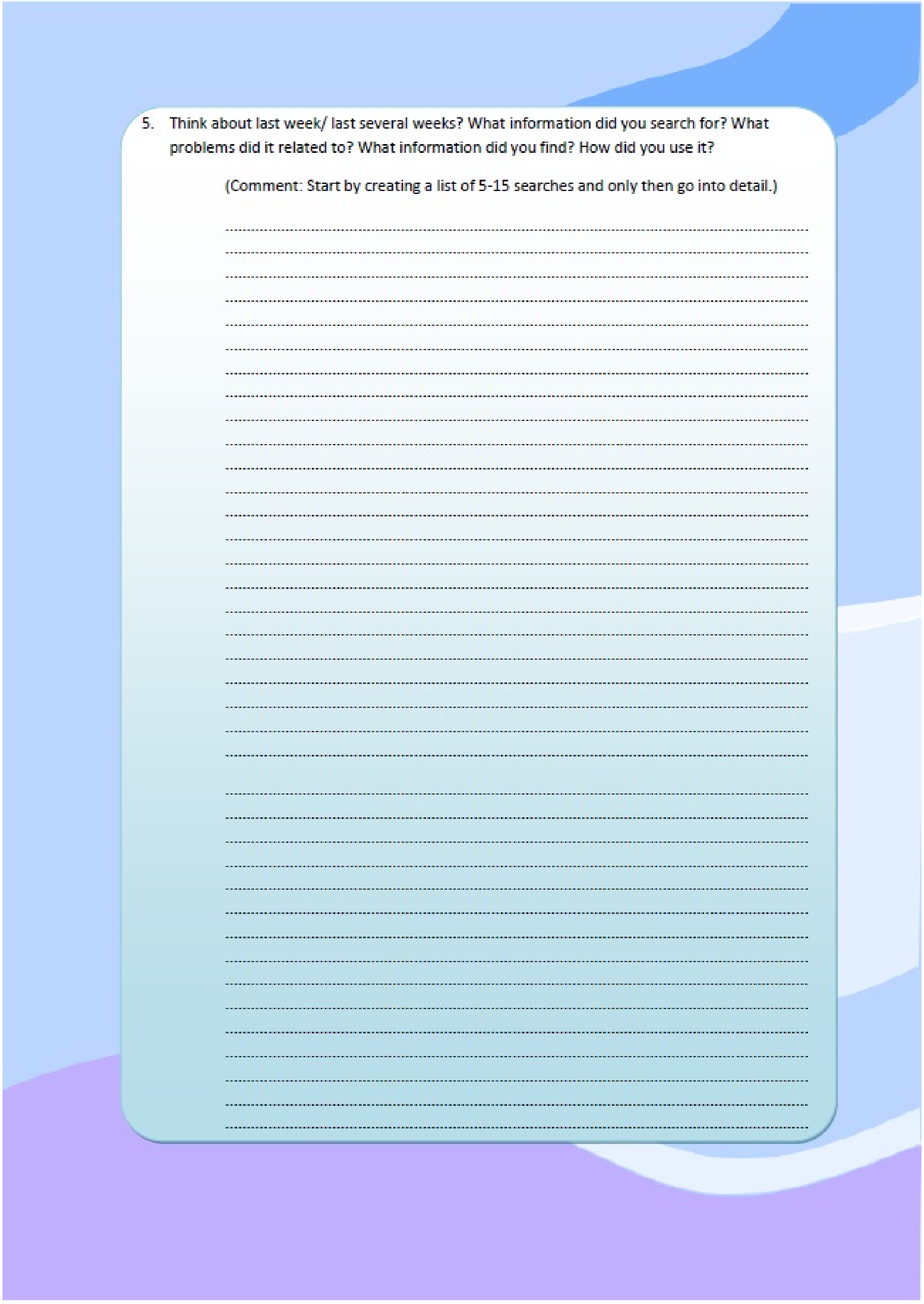
3. What are the key differences between working in the workplace and outside of it (focus on task distribution, information access and use and information needs)?

....

4. What are the key factors affecting your work related internet use?

....

Figure B.7: Interview Schedule 3 - Page 3 of the User Study 2 interview schedule



5. Think about last week/ last several weeks? What information did you search for? What problems did it related to? What information did you find? How did you use it?

(Comment: Start by creating a list of 5-15 searches and only then go into detail.)

Figure B.8: Interview Schedule 4 - Page 4 of the User Study 2 interview schedule

6. (For each of the listed searches) Please describe step by step the search process.

(Comment:

- Interface used – its advantages and shortcomings for this particular problem.
- Information object – its nature and the flow to other systems/people, factors which are used for relevance assessment
- Relevance related behaviour – what happens to the information once it is identified)

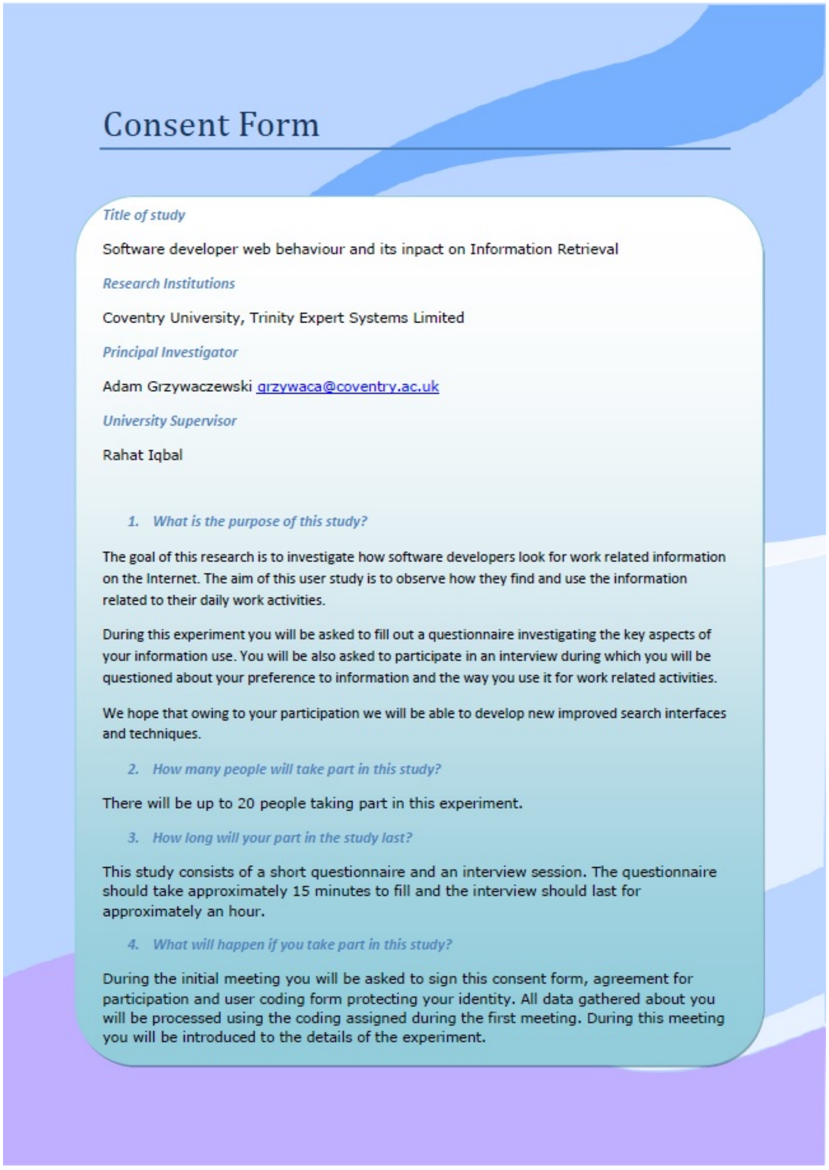
(Comment 2: During the discussion above ask about:

- Attitudes towards the behaviours
- Subjective norms
- Perception of control)

Figure B.9: Interview Schedule 5 - Page 5 of the User Study 2 interview schedule

B.4 Consent Form

User Study 2 consent form had a dual purpose. First of all it was a formal requirement of the Coventry University ethical approval process. Secondly it was key document that was used to explain the experimental procedure to the participants.



Consent Form

Title of study

Software developer web behaviour and its inpact on Information Retrieval

Research Institutions

Coventry University, Trinity Expert Systems Limited

Principal Investigator

Adam Grzywaczewski grzywaca@coventry.ac.uk

University Supervisor

Rahat Iqbal

1. *What is the purpose of this study?*

The goal of this research is to investigate how software developers look for work related information on the Internet. The aim of this user study is to observe how they find and use the information related to their daily work activities.

During this experiment you will be asked to fill out a questionnaire investigating the key aspects of your information use. You will be also asked to participate in an interview during which you will be questioned about your preference to information and the way you use it for work related activities.

We hope that owing to your participation we will be able to develop new improved search interfaces and techniques.

2. *How many people will take part in this study?*

There will be up to 20 people taking part in this experiment.

3. *How long will your part in the study last?*

This study consists of a short questionnaire and an interview session. The questionnaire should take approximately 15 minutes to fill and the interview should last for approximately an hour.

4. *What will happen if you take part in this study?*

During the initial meeting you will be asked to sign this consent form, agreement for participation and user coding form protecting your identity. All data gathered about you will be processed using the coding assigned during the first meeting. During this meeting you will be introduced to the details of the experiment.

Figure B.10: Consent Form 1 - Page 1 of the User Study 2 consent form

Following the meeting you will be asked to fill a short questionnaire that will be distributed to you over email. This will be followed by an interview session which will be recorded for the note taking purposes. The recordings will be destroyed after they were used for note taking and your identity will not be kept together with the notes protecting your identity.

5. *What are the potential benefits from being in this study?*

The ambition of this study is to identify software development related behaviours that can lead to development of new software development focused Information Retrieval utilities.

6. *What are the possible risks or discomforts involved from being in this study?*

The experiment does not introduce any risks to the participants.

7. *How will your privacy be protected?*

The identity of the participants will be gathered for the registration issues only. The identity of participants will not be disclosed neither used in the study. You will be assigned a code number and identified using the code number throughout the entire research.

8. *Will you receive anything for being in this study?*

No, the participation in the study is voluntary.

9. *Will it cost you anything to be in this study?*

There will be no costs involved.

10. *What if you have questions about this study?*

You can contact the leading researcher listed on the top of the document or his academic supervisor.

11. *What if you have questions about your rights as a research participant?*

In case of any queries or complains you can contact the Ethics Committee of Coventry University.

Figure B.11: Consent Form 2 - Page 2 of the User Study 2 consent form

In case of any queries or complains you can contact the Ethics Committee of Coventry University.

Participant's Agreement:

I have read the information provided above. I have asked all the questions I have at this time. I voluntarily agree to participate in this research study.

______________________________ ______________

Signature of Research Participant Date

____

Printed Name of Research Participant

______________________________ ______________

Signature of Person Obtaining Consent Date

____

Printed Name of Person Obtaining Consent

Figure B.12: Consent Form 3 - Page 3 of the User Study 2 consent form

Appendix C

User Study 3 Materials

C.1 Introduction

This appendix presents the key research materials used during the User Study 2 as well as very high level architectural overview of the developed experimental software. The appendix provides copies of the:

- Screening questionnaire
- Experiment description and consent documentation
- Overview of the experimental system architecture

C.2 Screening Questionnaire

The nature of software development processes and in more detail its organisation posed a real life challenge to the project. Software developers within the collaborating company engaged in a wide variety of activities (as expected based on the literature review) and proportion of them rarely delivered any actual source code. The developers that did develop code were engaged in the development of multiple distinct technologies, and used various development utilities, and quite frequently distinct development languages (C#, JavaScript, SQL variants and occasionally other) as well as a wide range of other content description languages (such as multiple variants of XML or HTML). To make matters even more complex developers rarely developed directly on their own machines but took extensive advantage of virtualisation (having multiple virtual machines for different projects and technologies) and more and more frequently working directly on remote servers hosting their environments (due to significant requirements of the technology based on which they were working). As a consequence it became apparent that it would not be possible to automatically observe the behaviour of all of the employees as for example it would not always be possible to install the observational software on a significant amount of virtual machines or give them access to the network. Similarly it was not practical to try to develop observation extensions for all possible development tools and browsers.

The above led to the development of the screening questionnaire illustrated below.

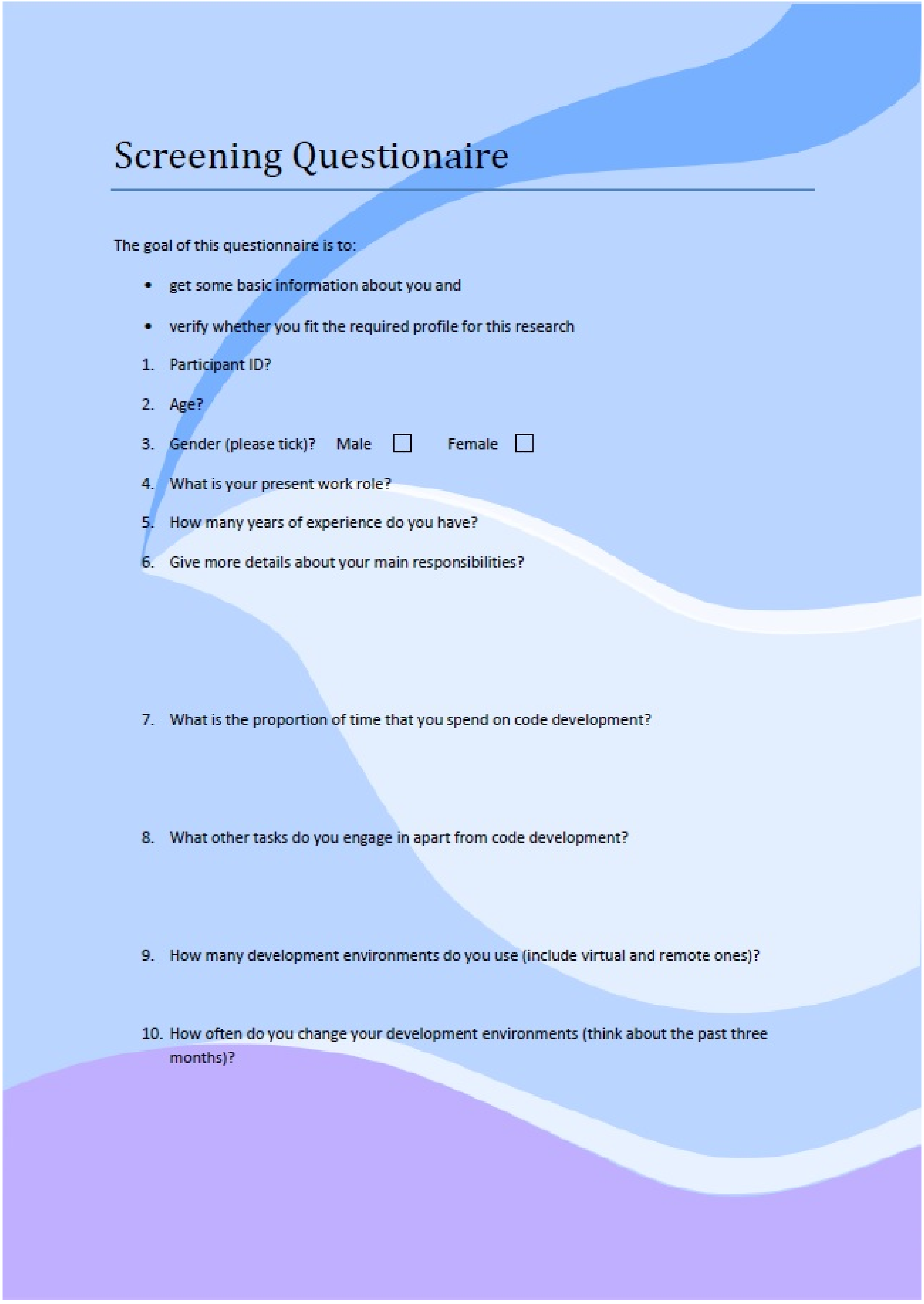
C.3 Consent Documentation

Since User Study 3 was not obstructive and did not require any active participation from the users it was essential to provide the experiment participants with a level of comfort and control. In order to achieve that, quite extensive consent documentation and experiment description was created.

The documentation not only goes into a lot of detail about how the information would be captured, pre-processed, transported and stored but also discussed exactly how it would be analysed and what would and would not be done with the data.

It was also essential to make sure that the experiment participants understood and accepted the level of engagement in the process. Even though, as mentioned earlier, the experiment did not require their intervention it was highly likely that the software would require maintenance during the experimentation (owing to its limited quality caused by the limited amount of development / test time that could be invested).

Similarly to User Study 1 the participation in the experiment was compensated so the documentation also formed a binding agreement and was a part of the financial process with the participant.



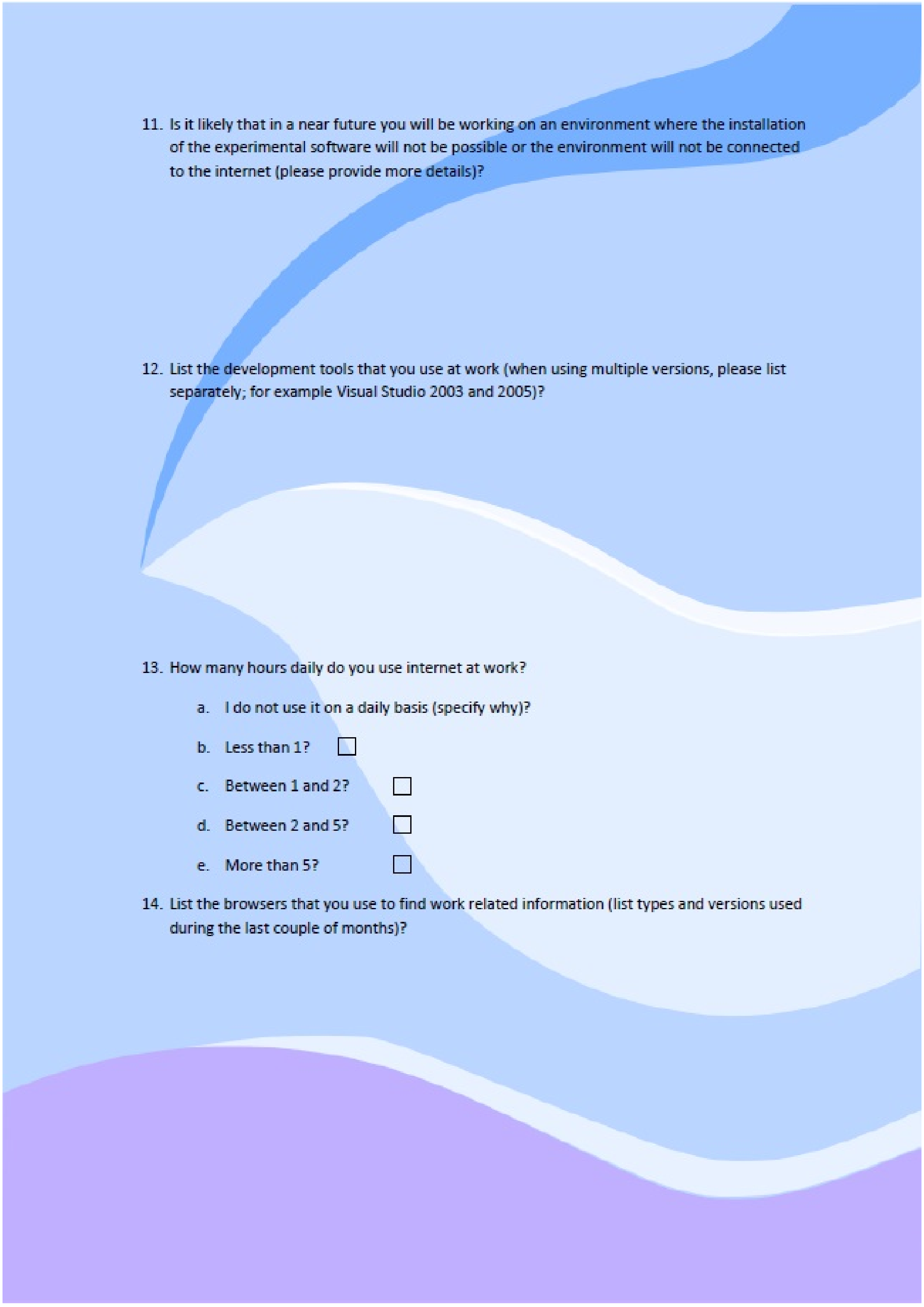
Screening Questionaire

The goal of this questionnaire is to:

- get some basic information about you and
- verify whether you fit the required profile for this research

1. Participant ID?
2. Age?
3. Gender (please tick)? Male ☐ Female ☐
4. What is your present work role?
5. How many years of experience do you have?
6. Give more details about your main responsibilities?
7. What is the proportion of time that you spend on code development?
8. What other tasks do you engage in apart from code development?
9. How many development environments do you use (include virtual and remote ones)?
10. How often do you change your development environments (think about the past three months)?

Figure C.1: Screening Questionnaire 1 - Page 1 of the User Study 3 screening questionnaire



11. Is it likely that in a near future you will be working on an environment where the installation of the experimental software will not be possible or the environment will not be connected to the internet (please provide more details)?

12. List the development tools that you use at work (when using multiple versions, please list separately; for example Visual Studio 2003 and 2005)?

13. How many hours daily do you use internet at work?
    a. I do not use it on a daily basis (specify why)?
    b. Less than 1? ☐
    c. Between 1 and 2? ☐
    d. Between 2 and 5? ☐
    e. More than 5? ☐

14. List the browsers that you use to find work related information (list types and versions used during the last couple of months)?

Figure C.2: Screening Questionnaire 2 - Page 2 of the User Study 3 screening questionnaire

15. List the browsers that you use for development and testing (list types and versions used during the last couple of months?

Figure C.3: Screening Questionnaire 3 - Page 3 of the User Study 3 screening questionnaire

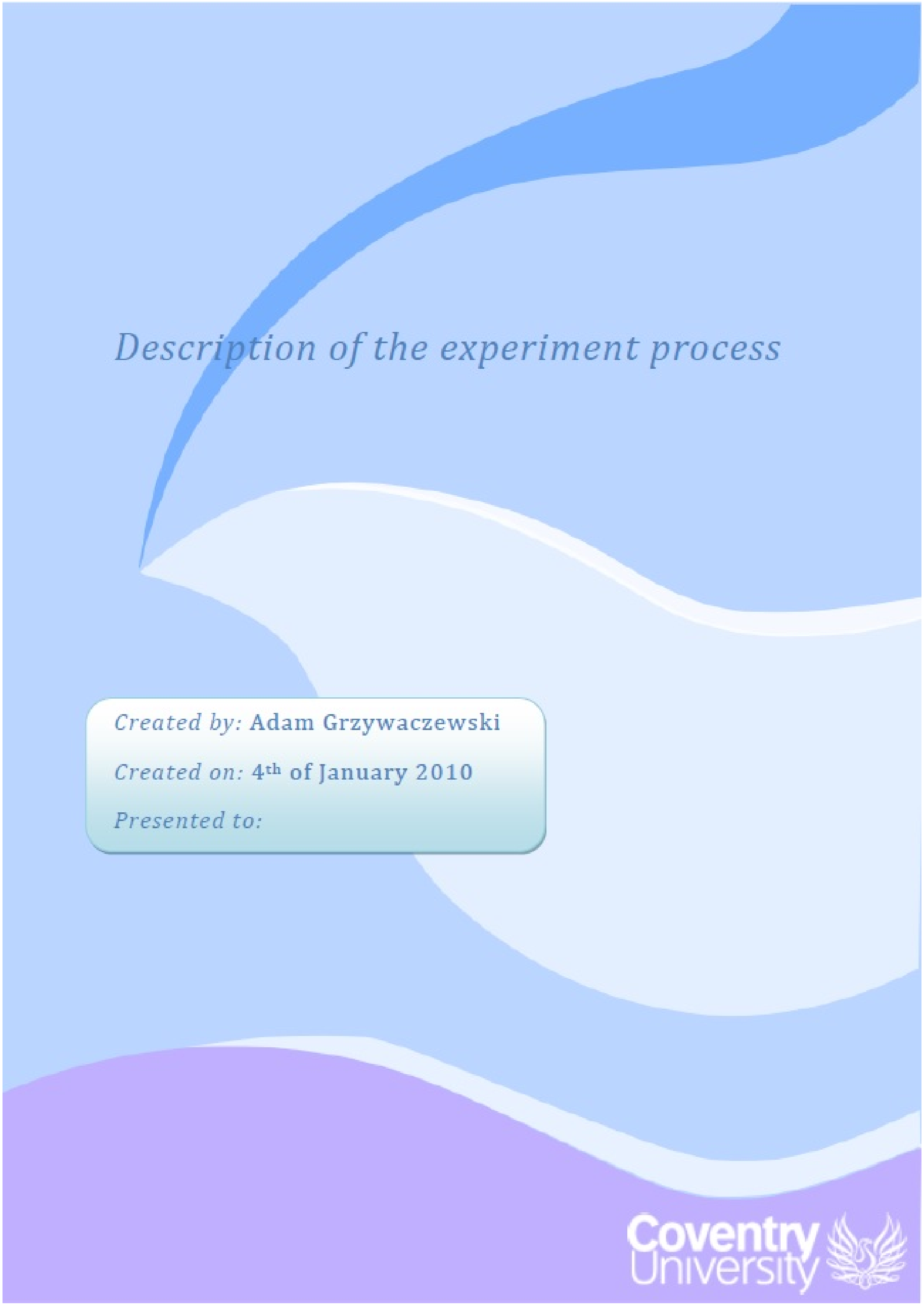


Figure C.4: Consent Documentation 1 - Page 1 of the User Study 3 consent documentation

1. The purpose of the document

The purpose of the document is to give the reader an overview of the third experiment performed for the research of Adam Grzywaczewski. This document will give the reader an overview of the problem, present the background of this experiment, present the goals of the experiment, its process and the way the captured data will be processed and analysed.

If you have any doubts or questions feel free to contact:

Adam Grzywaczewski	grzywaca@coventry.ac.uk	AS120b
Rahat Iqbal	R.Iqbal@coventry.ac.uk	AS218

Best Regards

The research team

Coventry University

Figure C.5: Consent Documentation 2 - Page 2 of the User Study 3 consent documentation

2. Research overview

The goal of this research is to investigate whether it is possible to automatically distinguish various tasks the user is performing basing on the automated observation of their browser interactions. This experiment is the third experiment of this particular research and aims at verifying whether the results obtained by automatic analysis of software developer behaviour can lead to identification of the user and the task the user is performing. We hope that the results of this experiment will help us to understand more the nature of human internet interactions and will lead to the improvement of the modern information retrieval systems.

Coventry University

Figure C.6: Consent Documentation 3 - Page 3 of the User Study 3 consent documentation

3. Goals of the experiment

The goal of the study is to perform a wide and long lasting study aiming to identify the behaviour of software developers during the work related activities. During the experiment the participants' interaction with the Internet Explorer, Firefox and Visual Studio will be observed and statistics of the interaction would be anonymously analysed. No content read or entered by the participant is being captured during the experimental process and the experiment focuses exclusively on statistics such as 'time the user spend on the page' (a full list of observed parameters can be found below). The data is fully anonymous and for the purpose of the analysis IDs' of the users will be considered exclusively.

Once the data is captured it will be analysed in search for traces of user consistency during the interaction with the computer. One of the main objectives is to determine whether for example the reading time is on average consistent for particular users and what are the differences between the users as far as the reading time is concerned. Additional statistics will be captured and refer to the section called Data Analysis for a full list of tests performed.

The experimental software can be turned on and off at any time by using the UI present in the tray but participants that will have the software running for more than 50% of their working time will be given an 8Gg memory stick and rewarded an additional £50.

Coventry University

Figure C.7: Consent Documentation 4 - Page 4 of the User Study 3 consent documentation

4. Experimental procedure

The experiment is composed of two stages:

- Questionnaire
- Automatic observation of behaviour

4.1. Opening questionnaire

The goal of the questionnaire is twofold:

- To determine the participants' suitability for the experiment (this is based on the users' choice of development tools, browser and pattern of work as well as basic aspects of user profile)
- To determine the basic characteristics of the user such as technology experience, etc.

4.2. Automated observation of behaviour

During the second stage the installed software captures user behaviour without the need for user intervention (unless new environment is created and installation is needed or update to the software is necessary).

The data captured by the software installed on participants machines will be transported over the secure channel to the university server. At the server the behavioural information will be stored and aggregates of information will be created.

Data will be collected from three sources: Internet Explorer, Firefox and Visual Studio (2005, 2008, 2010).

Data captured in the browser:

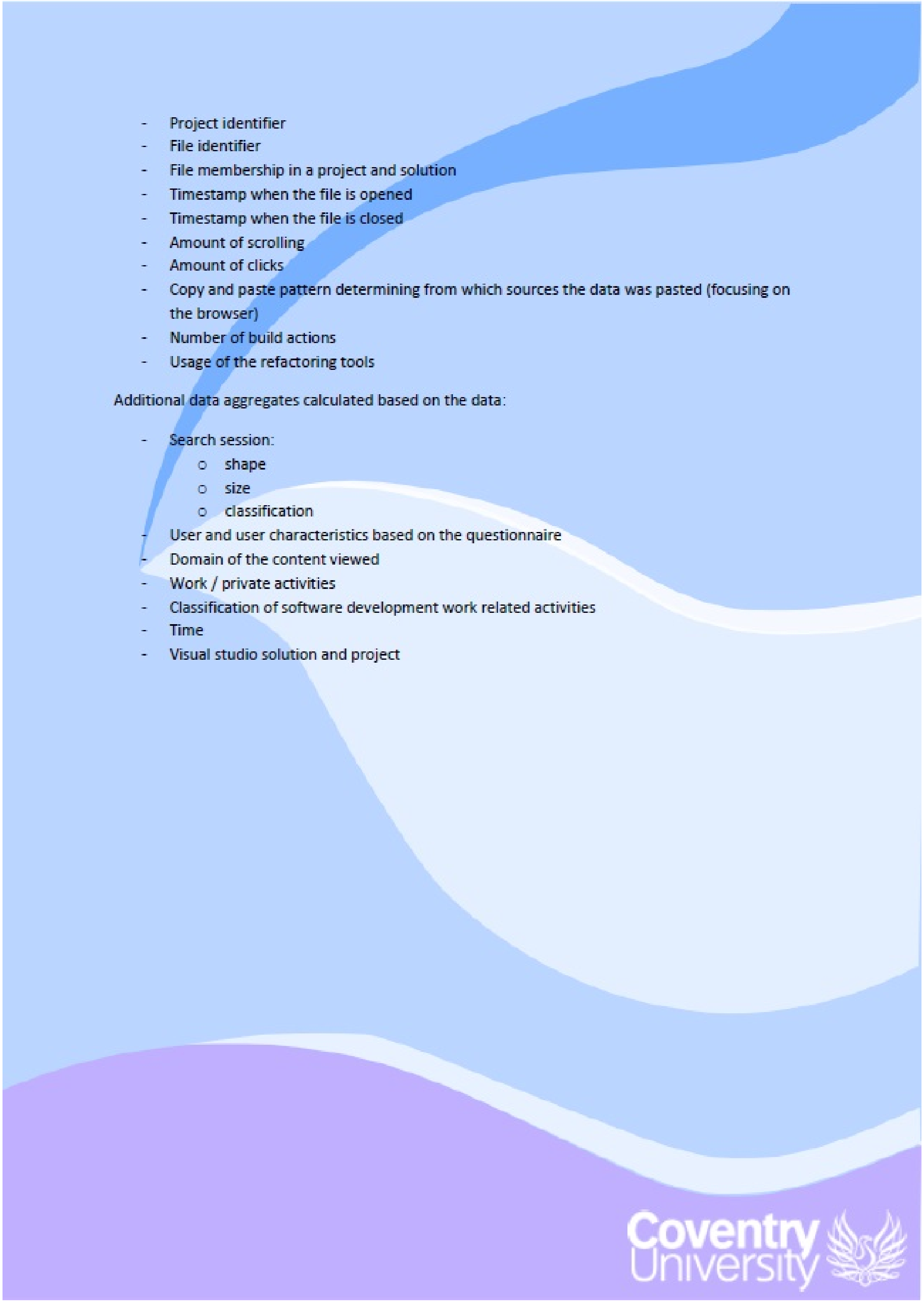
- Timestamp when the page is opened
- Timestamp when the page is closed
- Active reading time of the page
- Amount of scrolling
- Amount of clicks on the page
- Browser type
- Domain name of the page for classification of content as work / non work related
- GUID of the interaction
- GUID of the previous interaction (allowing to analyse the entire session and not only the individual pages)
- Amount of focus changes

Data captured in Visual Studio:

- Solution identifier

Coventry University

Figure C.8: Consent Documentation 5 - Page 5 of the User Study 3 consent documentation



- Project identifier
- File identifier
- File membership in a project and solution
- Timestamp when the file is opened
- Timestamp when the file is closed
- Amount of scrolling
- Amount of clicks
- Copy and paste pattern determining from which sources the data was pasted (focusing on the browser)
- Number of build actions
- Usage of the refactoring tools

Additional data aggregates calculated based on the data:

- Search session:
  - shape
  - size
  - classification
- User and user characteristics based on the questionnaire
- Domain of the content viewed
- Work / private activities
- Classification of software development work related activities
- Time
- Visual studio solution and project

Figure C.9: Consent Documentation 6 - Page 6 of the User Study 3 consent documentation

5. Experiment duration and participant engagement

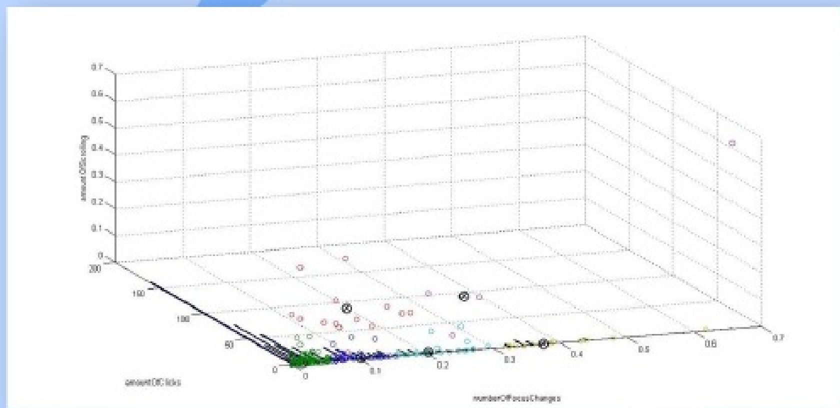
This is a third stage of the entire experimental process and its goal is to focus on long lasting pattern of interaction. The data will be analysed at different stages of the experiment but it is desired that the data will be captured for 2 to 6 months depending on the availability of participants. Overall though the engagement of participants will be minimal and should total to:

- Questioner 15 minutes
- Software installation 15 minutes
- Software support overall up to 30 minutes
- Software removal 5 minutes

Figure C.10: Consent Documentation 7 - Page 7 of the User Study 3 consent documentation

6. Data processing and analysis

The main focus of the data analysis is to determine the stability of the distribution of behaviour across various users, parameters of the session and time. Let's consider the following distribution of data:



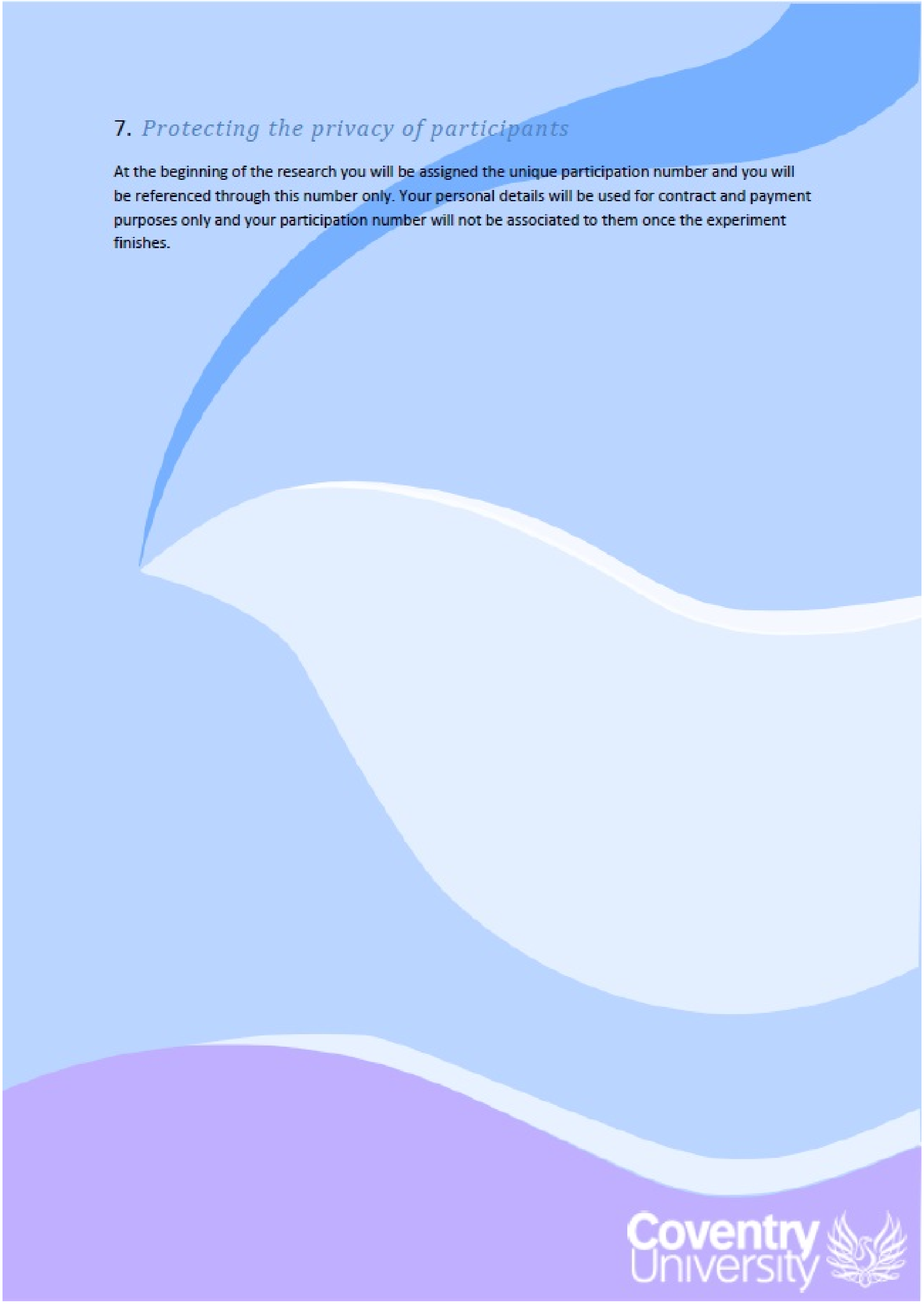
The single points on the image above represent the particular interactions of the user with the web browser and the histogram plotted on one of the axes represents the number of points that fall into the defined intervals.

On high level the goal of data analysis is to determine:

1. What is the distribution of data represented by the histogram
2. How the histogram changes for particular users in time
3. How the histogram changes across various users and their different characteristics such as experience
4. How the histogram changes across work/non work related tasks
5. How the histogram changes across different types of software development work related tasks
6. How the histogram changes across different VS solutions and projects
7. How the histogram changes across different web browser sessions
8. Whether the differences described above are statistically significant and what are their relative magnitudes

Coventry University

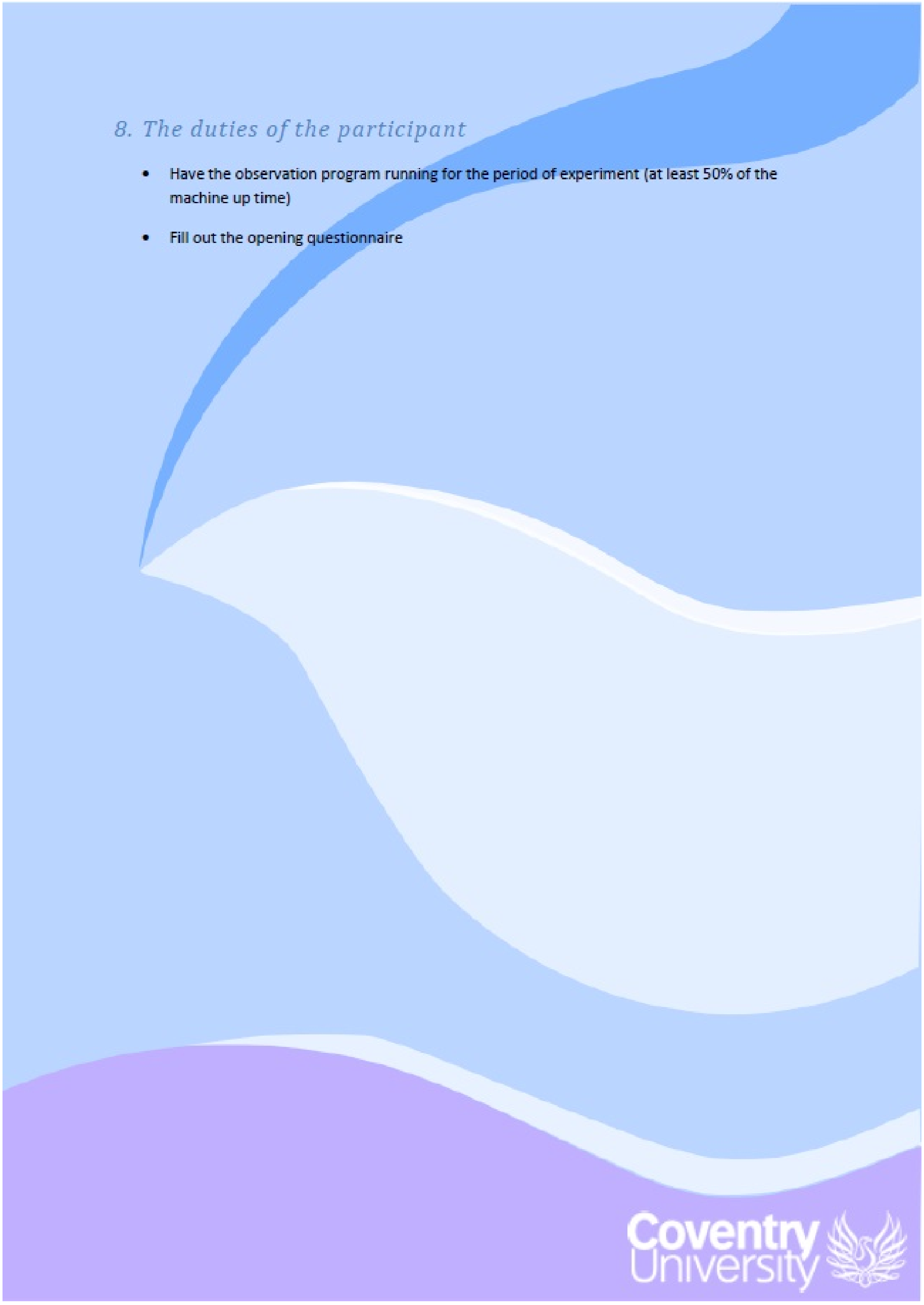
Figure C.11: Consent Documentation 8 - Page 8 of the User Study 3 consent documentation



7. *Protecting the privacy of participants*

At the beginning of the research you will be assigned the unique participation number and you will be referenced through this number only. Your personal details will be used for contract and payment purposes only and your participation number will not be associated to them once the experiment finishes.

Figure C.12: Consent Documentation 9 - Page 9 of the User Study 3 consent documentation



8. The duties of the participant

- Have the observation program running for the period of experiment (at least 50% of the machine up time)
- Fill out the opening questionnaire

Figure C.13: Consent Documentation 10 - Page 10 of the User Study 3 consent documentation

Consent Form

Title of study

Software developer web behaviour and its inpact on Information Retrieval

Research Institutions

Coventry University, Trinity Expert Systems Limited

Principal Investigator

Adam Grzywaczewski grzywaca@coventry.ac.uk

University Supervisor

Rahat Iqbal

1. *What is the purpose of this study?*

The goal of this research is to investigate how software developers look for work related information on the Internet. The aim of this user study is to observe how they choose the search strategies and to what extent they are consistent in choosing them.

During this experiment you will be asked to fill out a questionnaire describing your main characteristics such as your name, the browsers you use, etc. Secondly you will be asked to install the experimental software on your machine for the period of 2-6 months that will capture statistics of your behaviour such as average reading time, average number of clicks, web session size and shape, etc.

We hope that owing to your participation we will be able to develop new improved search interfaces and techniques.

2. *How many people will take part in this study?*

There will be up to 20 people taking part in this experiment.

3. *How long will your part in the study last?*

For the purpose of this study you will be asked to install experimental software that will be capturing your daily internet and visual studio activities. The software will not consume your time in any way and you will not be asked to interact with it in any way. You will be asked though to install the updates of the software occasionally and report any failures of the software. Apart from that you will be asked to fill out a questionnaire that will capture your basic characteristics such as age, gender, software development experience, browsers used, etc.

Figure C.14: Consent Documentation 11 - Page 11 of the User Study 3 consent documentation

4. *What will happen if you take part in this study?*

During the initial meeting you will be asked to sign this consent form, agreement for participation and user coding form protecting your identity. All data gathered about you will be processed using the coding assigned during the first meeting. During this meeting you will be introduced to the details of the experiment.

During the course of the study you will have software installed on your machine gathering measurements of your behaviour such as mouse click frequencies and number of visits on sites.

5. *What are the potential benefits from being in this study?*

Hopefully the results of this study will help to identify a generic and topic independent way of search task classification. Hopefully this research will help to understand the potential of personalized and context aware search and in the long term improve the existing search algorithms.

6. *What are the possible risks or discomforts involved from being in this study?*

The experiment does not introduce any risks to the participants.

7. *How will your privacy be protected?*

The identity of the participants will be gathered for the registration issues only. The identity of participants will not be disclosed neither used in the study. You will be assigned a code number and identified using the code number throughout the entire research.

8. *Will you receive anything for being in this study?*

Yes, participants that will have the experimental software running for at least 50% of their machine up time and will take part in the experiment for at least 2 months will receive £50 and an 8Gb memory stick.

9. *Will it cost you anything to be in this study?*

There will be no costs involved.

10. *What if you have questions about this study?*

You can contact the leading researcher listed on the top of the document or his academic supervisor.

Figure C.15: Consent Documentation 12 - Page 12 of the User Study 3 consent documentation

11. What if you have questions about your rights as a research participant?

In case of any queries or complains you can contact the Ethics Committee of Coventry University.

Participant's Agreement:

I have read the information provided above. I have asked all the questions I have at this time. I voluntarily agree to participate in this research study.

________________________________ ______________

Signature of Research Participant Date

____

Printed Name of Research Participant

________________________________ ______________

Signature of Person Obtaining Consent Date

____

Printed Name of Person Obtaining Consent

Figure C.16: Consent Documentation 13 - Page 13 of the User Study 3 consent documentation

C.4 System Architecture

The final section of this appendix illustrates the architecture of the solution utilised for the purpose of experimentation.

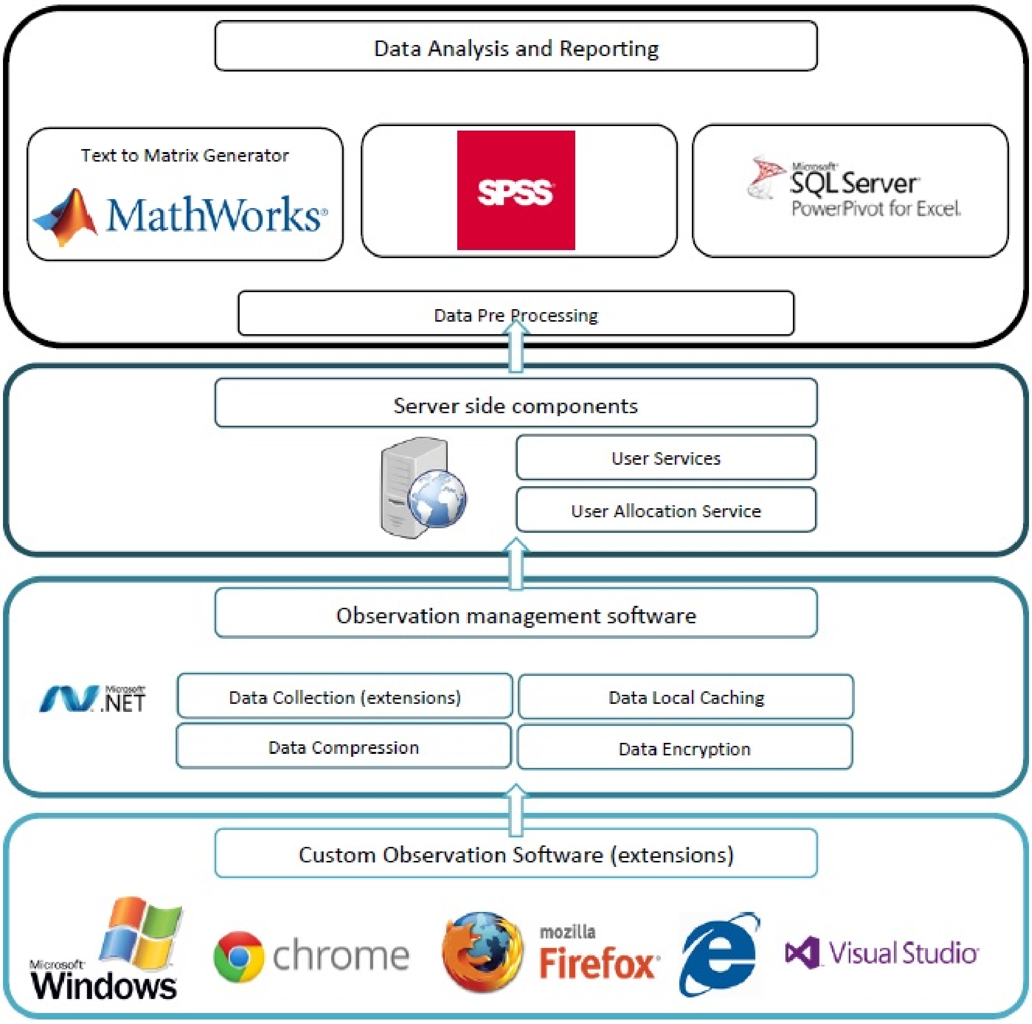


Figure C.17: Experimental System Architecture - The graphical illustration of the architecture of the software utilised during User Study 3

Publications

Grzywaczewski, A. and Iqbal, R. (2011), Software engineers' information behaviour and implicit relevance indicators., pp. 185–201.

Grzywaczewski, A. and Iqbal, R. (2012), Task-specific information retrieval systems for software engineers., pp. 1204–1218.

Grzywaczewski, A., Iqbal, R., James, A. and Halloran, J. (2009*a*), Developing a context-aware information retrieval system, *in* '4th International Workshop on Ubiquitous Computing', pp. 37–54.

Grzywaczewski, A., Iqbal, R., James, A. and Halloran, J. (2011*a*), *Emerging Pervasive and Ubiquitous Aspects of Information Systems: Cross-Disciplinary Advancements*, Premier Reference Source, Igi Global, chapter An Approach for Capturing Human Information Behaviour.
URL: *http://books.google.co.uk/books?id=QqpAYqpjFbMC*

Grzywaczewski, A., Iqbal, R., James, A. and Halloran, J. (2011*b*), Software developers' information needs: towards the development of intelligent recommender systems, *in* 'Proceedings of the 6th international conference on Ubiquitous and Collaborative Computing', iUBICOM'11, British Computer Society, Swinton, UK, UK, pp. 66–74.
URL: *http://dl.acm.org/citation.cfm?id=2227369.2227377*

Grzywaczewski, A., Iqbal, R., James, A. and Halloran, J. (2012), Copy and paste as an indicator of relevance: Towards the development of intelligent recommender systems for software engineers, *in* 'Task-Based and Aggregated Search (TBAS2012) Workshop', ECIR 2012, pp. 14–18.

Grzywaczewski, A., Iqbal, R., James, R. and Halloran, J. (2009*b*), An investigation of user behaviour consistency for context-aware information retrieval systems., pp. 69–90.

Grzywaczewski, A., Iqbal, R., Shah, N. and James, A. (2010), E-marketing strategy for businesses, *in* 'e-Business Engineering (ICEBE), 2010 IEEE 7th International Conference on', pp. 428 –434.

Halloran, J., Iqbal, R., Aliakseyeu, D., Fernando, M., Cooper, R., Grzywaczewski, A., Grewal, R., James, A. and Greenhalgh, C. (2009), Design challenges and solutions: Review of the 4th international workshop on ubiquitous computing (iubicom 2009), *in* 'Proceedings of the 26th British National Conference on Databases: Dataspace: The Final Frontier', BNCOD 26, Springer-Verlag, Berlin, Heidelberg, pp. 234–245.

Iqbal, R., Grzywaczewski, A., James, A. E., Doctor, F. and Halloran, J. (2012), Investigating the value of retention actions as a source of relevance information in the software development environment., *in* 'CSCWD', pp. 121–127.

References

Aarts, H. and Dijksterhuis, A. (2000), 'Habits as knowledge structures: automaticity in goal directed behavior.', *Journal of personality and social psychology* **78**(1), 53–63. 101, 123, 77

Abelson, R. (1968), *Theories of cognitive consistency; a sourcebook*, Rand McNally. 38, 32

Ajzen, I. (1991), 'The theory of planned behavior', *Organizational Behavior and Human Decision Processes* **50**(2), 179–211. 97, 101, 102

Ajzen, I. (2002), 'Residual effects of past on later behavior: Habituation and reasoned action perspectives', *Personality and Social Psychology Review* **6**(2), 107–122. 15, 101, 114, 123, 7, 72, 77

Ajzen, I. (2005), *Attitudes, Personality and Behavior*, Mapping social pschology, McGraw-Hill Education. 6, 9, 14, 26, 37, 38, 64, 97, 132, 2, 7, 20, 31, 32, 46, 83

Ajzen, I. (2011), 'The theory of planned behaviour: Reactions and reflections', *Psychology and Health* **26**(9), 1113–1127. 14, 97, 98, 101, 103, 7

Ajzen, I. and Fishbein, M. (1977), 'Attitude-behavior relations: A theoretical analysis and review of empirical research.', *Psychological bulletin* **84**(5), 888. 35, 37, 38, 29, 31, 32

Albarracn, D., Johnson, B. and Zanna, M. (2005), *The handbook of attitudes*, Lawrence Erlbaum. 63, 64, 45

Allen, J. F. (1984), 'Towards a general theory of action and time', *Artif. Intell.* **23**(2), 123–154.
URL: *http://dx.doi.org/10.1016/0004-3702(84)90008-0* 44

Almutairi, H. (2011), 'Factors affecting the information behaviour of managers in the kuwaiti civil service: a relational model', *Information Research* **16**(2). 4, 72, 54

Anderson, J. (1982), 'Acquisition of cognitive skill.', *Psychological review* **89**(4), 369. 68, 50

Anderson, J., Bothell, D., Byrne, M., Douglass, S., Lebiere, C. and Qin, Y. (2004), 'An integrated theory of the mind.', *Psychological review* **111**(4), 1036. 68, 50

Anderson, J., Boyle, C. and Yost, G. (1985), The geometry tutor, *in* 'Byte', Vol. 1, pp. 1–7. 68, 50

Armitage, C. and Conner, M. (2001), 'Efficacy of the theory of planned behaviour: A meta-analytic review', *British Journal of Social Psychology* **40**(4), 471–499. 6, 14, 97, 103, 7

Arvola, A., Vassallo, M., Dean, M., Lampila, P., Saba, A., Lähteenmäki, L. and Shepherd, R. (2008), 'Predicting intentions to purchase organic food: The role of affective and moral attitudes in the theory of planned behaviour', *Appetite* **50**(2), 443–454. 101

Back, J. and Oppenheim, C. (2001), 'A model of cognitive load for IR: implications for user relevance feedback interaction', *Information Research* **6**(2). 10, 3

Bae, H.-S. (2008), 'Entertainment-education and recruitment of cornea donors: The role of emotion and issue involvement', *Journal of Health Communication* **13**(1), 20–36. 101

Bar-Ilan, J., Keenoy, K., Levene, M. and Yaari, E. (2009), 'Presentation bias is significant in determining user preference for search results:a user study', *Journal of the American Society for Information Science and Technology* **60**(1), 135–149. 3, 4, 12, 5

Bargh, J. (2007), *Social Psychology And the Unconscious: The Automaticity of Higher Mental Processes*, Frontiers of Social Psychology Series, Taylor & Francis Group. 15, 114, 7, 72

Bates, M. J. (1989), 'The design of browsing and berrypicking techniques for the online search interface', *Online Information Review* **13**(5), 407–424. 52, 54, 55

Batra, R. and Ahtola, O. (1991), 'Measuring the hedonic and utilitarian sources of consumer attitudes', *Marketing Letters* **2**(2), 159–170. 101

Belkin, N. (1980), 'Anomalous states of knowledge as a basis for information retrieval', *Canadian journal of information science* **5**(1), 133–143. 4, 44, 45, 46

Belkin, N. J. (2008), 'Some(what) grand challenges for information retrieval.', *SIGIR Forum* **42**(1), 47–54. 5

Belkin, N. J., Oddy, R. N. and Brooks, H. M. (1982), 'Ask for information retrieval: Part i. background and theory', *Journal of documentation* **38**(2), 61–71. 46

Bem, D. (1967), 'Self-perception: An alternative interpretation of cognitive dissonance phenomena.', *Psychological review* **74**(3), 183. 44

Bichteler, J. and Ward, D. (1989), 'Information-seeking behavior of geoscientists.', *Special Libraries* **80**(3), 169–78. 117

Brandt, J., Dontcheva, M., Weskamp, M. and Klemmer, S. (2010*a*), Example-centric programming: integrating web search into the development environment, *in* 'Proceedings of the 28th international conference on Human factors in computing systems', ACM, pp. 513–522. 134, 143, 84, 92

Brandt, J., Dontcheva, M., Weskamp, M. and Klemmer, S. (2010*b*), Example centric programming: integrating web search into the development environment, *in* 'Proceedings of the 28th international conference on Human factors in computing systems', ACM, New York, NY, USA, pp. 513–522. 146, 84, 95

REFERENCES

Brandt, J., Guo, P., Lewenstein, J., Dontcheva, M. and Klemmer, S. (2009), Two studies of opportunistic programming: interleaving web foraging, learning, and writing code, *in* 'Proceedings of the SIGCHI Conference on Human Factors in Computing Systems', CHI '09, ACM, New York, NY, USA, pp. 1589–1598. 121, 122, 143, 185, 74, 76, 84, 92

Brandt, J., Guo, P., Lewenstein, J. and Klemmer, S. (2008), Opportunistic programming: how rapid ideation and prototyping occur in practice, *in* 'WEUSE '08: Proceedings of the 4th international workshop on End-user software engineering', ACM, New York, NY, USA, pp. 1–5. 143, 84, 92

Brown, C. M. (1999), 'Information seeking behavior of scientists in the electronic information age: Astronomers, chemists, mathematicians, and physicists', *JASIS* **50**(10), 929–943. 117

Bruce, H., Fidel, R., Pejtersen, A. M., Dumais, S., Grudin, J. and Poltrock, S. (2003), 'A comparison of the collaborative information retrieval behaviour of two design teams', *The New Review of Information Behaviour Research* **4**(1), 139–153. 117

Buckley, J., O'Brien, M. and Power, N. (2006), Empirically refining a model of programmers' information-seeking behavior during software maintenance, *in* '18th Annual Psychology of Programming Interest Group (PPIG) Workshop'. 131, 133, 136, 137, 81, 84

Buda, R. and Zhang, Y. (2000), 'Consumer product evaluation: the interactive effect of message framing, presentation order, and source credibility', *Journal of Product and Brand Management* **9**(4), 229–242. 4, 52, 40

Burgin, R. (1992), 'Variations in relevance judgments and the evaluation of retrieval performance', *Information Processing and Management* **28**(5), 619–627. 30, 24

Bystrom, K. and Jarvelin, K. (1995), 'Task complexity affects information seeking and use', *Information processing & management* **31**(2), 191–213. 59

Campbell, D. (1988), 'Task complexity: A review and analysis', *Academy of management review* pp. 40–52. 59

Case, D. O. (1986), 'Collection and organization of written information by social scientists and humanists: a review and exploratory study', *Journal of Information Science* **12**(3), 97–104. 117, 119

Case, D. O. (2007), *Looking for information*, Emerald Group Pub Ltd. 4, 12, 14, 40, 198, 5, 6, 34, 140

Chomsky, N. (1959), 'A review of b.f. skinner's verbal behavior', *Language* **35**(1), 26–58. 73, 55

Choo, C. W. and Marton, C. (2003), 'Information seeking on the web by women in it professions', *Internet Research* **13**(4), 267–280. 136, 137

Cleverdon, C. W., Mills, J. and Keen, M. (1966), *Factors determining the performance of indexing systems*, ASLIB Cranfield project, Cranfield. 27, 21

Condry, J. (1977), 'Enemies of exploration: Self-initiated versus other-initiated learning.', *Journal of Personality and Social Psychology* **35**(7), 459. 44

Cool, C. (2001), 'The concept of situation in information science.', *Annual Review of Information Science and Technology American Society for Information Science* **35**(1), 5–42. 4

Cooper, J. and Croyle, R. (1984), 'Attitudes and attitude change', *Annual review of psychology* **35**(1), 395–426. 63, 45

Cooper, M. and Chen, H. (2001), 'Predicting the relevance of a library catalog search', *Journal of the American Society for Information Science and Technology* **52**(10), 813–827. 12, 5

Cuadra, C. and Katter, R. (1967), 'Opening the black box of 'relevance'', *Journal of Documentation* **23**(4), 291–303. 28, 22

Cuadra, C., Katter, R., Holmes, E., Wallace, E. and Monica, C. (1967), *Experimental Studies of Relevance Judgments: Final Report*, number v. 1 *in* 'Technical memorandum, TM series', System Development Corporation and National Science Foundation (U.S.). 28, 22

Curtis, B., Krasner, H. and Iscoe, N. (1988), 'A field study of the software design process for large systems', *Communications of the ACM* **31**(11), 1268–1287. 120, 131, 74, 82

Davenport, E. (2010), 'Confessional methods and everyday life information seeking', *Annual Review of Information Science and Technology* **44**(1), 533–562.
URL: *http://dx.doi.org/10.1002/aris.2010.1440440119* 43

Day, R. E. (2011), 'Death of the user: Reconceptualizing subjects, objects, and their relations', *JASIST* **62**(1), 78–88. 44

De Solla Price, D. J. (1965), 'Is technology historically independent of science? a study in statistical historiography', *Technology and Culture* pp. 553–568. 117

Dervin, B. (1983), *An overview of sense-making research: concepts, methods, and results to date*, Sense-making packet, The Author. 4, 15, 45, 8

Dervin, B. (1997), Given a context by any other name: Methodological tools for taming the unruly beast, *in* 'Information Seeking in Context', Taylor Graham, pp. 13–38. 4

Dervin, B. (1999), 'Chaos, order and sense-making: A proposed theory for information design', *Information design* pp. 35–57. 41

Dervin, B. and Clark, K. (1999), 'Exemplars of the use of the sense-making methodology (meta-theory and method): In-depth introduction to the sense-making issues of the electronic journal of communication', *Electronic Journal of Communication* **9**(2). 12, 5

Dervin, B., Fisher, K. E., Durrance, J., Ross, C., Savolainen, R. and Solomon, P. (2005), 'Reports of the demise of the "user" have been greatly exaggerated: Dervin's sense-making and the methodological resuscitation of the user looking backwards, looking forward', *Proceedings of the American Society for Information Science and Technology* **42**(1), n/a–n/a.
URL: *http://dx.doi.org/10.1002/meet.1450420159* 40

Dervin, B., Foreman-Wernet, L. and Lauterbach, E. (2003), *Sense-making methodology reader: Selected Writings of Brenda Dervin*, Hampton Press. 12, 14, 15, 40, 42, 45, 49, 5, 6, 8, 34, 35

Detienne, F. and Bott, F. (2001), *Software Design - Cognitive Aspects*, Practitioner Series, Springer. 115, 133, 138, 73, 84, 87

Détienne, F. and Bott, F. (2002), *Software design - cognitive aspects*, Springer Verlag. 133, 138, 84, 87

Dong, Y. and Lee, K. (2008), 'A cross-cultural comparative study of users perceptions of a webpage: With a focus on the cognitive styles of chinese, koreans and americans', *International Journal of Design* **2**(2), 19–30. 4

Duckitt, J. (1992), 'Prejudice and behavior: A review', *Current Psychology* **11**(4), 291–307. 35, 29

Duckitt, J. (1994), *The social psychology of prejudice*, Praeger Publishers. 35, 29

Eagly, A. and Chaiken, S. (1998), 'Attitude structure and function.'. xix, xv

Edwards, K. (1990), 'The interplay of affect and cognition in attitude formation and change.', *Journal of personality and social psychology* **59**(2), 202. 101

Ellis, D. (1996), 'The dilemma of measurement in information retrieval research', *Journal of the American Society for Information Science* **47**(1), 23–36. 9, 27, 2, 21

Ellis, D., Cox, D. and Hall, K. (1993), 'A comparison of the information seeking patterns of researchers in the physical and social sciences', *Journal of documentation* **49**(4), 356–369. 117

Ellis, D. and Haugan, M. (1997), 'Modelling the information seeking patterns of engineers and research scientists in an industrial environment', *Journal of Documentation* **53**(4), 384–403. 4, 131, 134, 136, 81, 84

Elsenberg, M. and Barr, C. (1988), 'Order effects: A study of the possible inuence of presentation order on user judgments of document relevance', *Journal of the American Society for Information Science and Technology* **39**, 293–301. 4, 12, 52, 5, 40

Engel, A., Herr, T., Johnson, K., Gick, M., Holyoak, K. and Woodward, E. (1999), *Problem-Solving Strategies*, Problem Books in Mathematics, Springer. 69, 51

Fazio, R. (1990), 'Multiple processes by which attitudes guide behavior: The mode model as an integrative framework', *Advances in experimental social psychology* **23**, 75–109. 63, 101, 45

Fazio, R. and Towles-Schwen, T. (1999), 'The mode model of attitude-behavior processes', *Dual process theories in social psychology* pp. 97–116. 39, 63, 34, 45

Fazio, R. and Zanna, M. (1981), 'Direct experience and attitude-behavior consistency', *Advances in experimental social psychology* **14**, 161–202. 63, 45

Fidel, R. and Green, M. (2004), 'The many faces of accessibility: engineers' perception of information sources', *Information Processing & Management* **40**(3), 563–581. 117, 119

Fidel, R., Pejtersen, A., Cleal, B. and Bruce, H. (2004), 'A multidimensional approach to the study of human-information interaction: A case study of collaborative information retrieval', *Journal of the American Society for Information Science and Technology* **55**(11), 939–953. 12, 14, 56, 57, 58, 5, 6, 43, 44

Fishbein, M. and Ajzen, I. (1975), *Belief, attitude, intention, and behavior*, Addison Wesley Publishing Company. 97

Fishbein, M. and Ajzen, I. (2009), *Predicting and Changing Behavior*, Psychology Press. 14, 101, 103, 133, 7, 83

Fisher, K. E., Unruh, K. T. and Durrance, J. C. (2003), 'Information communities: Characteristics gleaned from studies of three online networks', *Proceedings of the American Society for Information Science and Technology* **40**(1), 298–305. 41

Fisher, K. and Julien, H. (2005), *Theories of information behavior*, Vol. 43, Information Today Inc. 4, 12, 14, 40, 50, 55, 5, 7, 34, 37

Flaxbart, D. (2001), 'Conversations with chemists: Information-seeking behavior of chemistry faculty in the electronic age', *Science & Technology Libraries* **21**(3-4), 5–26. 117

Ford, N. (2004), 'Modeling cognitive processes in information seeking: From popper to pask', *Journal of the American Society for Information Science and Technology* **55**(9), 769–782. 4

Ford, N., Miller, D. and Moss, N. (2005), 'Web search strategies and human individual differences: Cognitive and demographic factors, internet attitudes, and approaches', *Journal of the American Society for Information Science and Technology* **56**(7), 741–756. 4, 72, 54

Ford, N., Wilson, T., Foster, A., Ellis, D. and Spink, A. (2002), 'Information seeking and mediated searching. part 4. cognitive styles in information seeking', *Journal of the American Society for Information Science and Technology* **53**(9), 728–735. 3

Forssell, D. and Powers, W. (2009), *Perceptual Control Theory*, Living Control Systems Pub. 72, 74, 75, 54, 56, 57

French, D. P., Sutton, S., Hennings, S. J., Mitchell, J., Wareham, N. J., Griffin, S., Hardeman, W. and Kinmonth, A. L. (2005), 'The importance of affective beliefs and attitudes in the theory of planned behavior: Predicting intention to increase physical activity1', *Journal of Applied Social Psychology* **35**(9), 1824–1848. 101

Freund, L., Toms, E. and Waterhouse, J. (2005), 'Modeling the information behaviour of software engineers using a work-task framework', *Proceedings of the American Society for Information Science and Technology* **42**(1). 115, 117, 121, 122, 124, 125, 131, 134, 138, 187, 73, 74, 76, 77, 78, 79, 81, 82, 84, 87, 129

Gallardo-Valencia, R. and Sim, S. (2009), Internet-scale code search, *in* 'ICSE Workshop on Search-Driven Development-Users, Infrastructure, Tools and Evaluation (SUITE '09)', IEEE Computer Society, Washington, DC, USA, pp. 49–52. 131, 134, 138, 82, 84, 87

REFERENCES

Gallardo-Valencia, R., Tantikul, P. and Sim, S. (2010), Searching for reputable source code on the web, *in* 'Proceedings of the 16th ACM international conference on Supporting group work', GROUP '10, ACM, New York, NY, USA, pp. 183–186. xvii, 131, 138, 139, xiii, 82, 87, 88

Gerstberger, P. G. and Allen, T. J. (1968), 'Criteria used by research and development engineers in the selection of an information source.', *Journal of applied psychology* **52**(4), 272. 117

Gick, M. (1986), 'Problem-solving strategies.', *Educational Psychologist* **21**, 99–120. 43, 36

Giddens, A. (1984), *The constitution of society: Outline of the theory of structuration*, Univ of California Press. 6

Goodmonson, C. and Glaudin, V. (1971), 'The relationship of commitment-free behavior and commitment behavior: A study of attitude toward organ transplantation', *Journal of Social Issues* **27**(4), 171–183. 63, 45

Graff, M., Davies, J. and McNorton, M. (2001), 'Cognitive style and cross cultural differences in internet use and computer attitudes', *European Journal of Open, Distance and E-Learning* **2**. 3, 4

Gralewska-Vickery, A. (1976), 'Communication and information needs of earth science engineers', *Information Processing & Management* **12**(4), 251–282. 117

Greeno, J. (1978), *Handbook of learning and cognitive processes: V. Human information*, L. Erlbaum Associates, chapter Natures of problem-solving abilities. 69, 51

Gross, M. (2005), 'The impact of low-level skills on information-seeking behavior: Implications of competency theory for research and practice', *Reference & User Services Quarterly* pp. 155–162. 4

Gull, C. (1956), 'Seven years of work on the organization of materials in the special library', *American Documentation* **7**(4), 320–329. 27, 21

Gutwin, C. and Greenberg, S. (1998), Design for individuals, design for groups: tradeoffs between power and workspace awareness, *in* 'Proceedings of the 1998 ACM conference on Computer supported cooperative work', CSCW '98, ACM, New York, NY, USA, pp. 207–216.
URL: *http://doi.acm.org/10.1145/289444.289495* 45

Guzdial, M., Ludovice, P., Realff, M., Morley, T. and Carroll, K. (2002), When collaboration doesnt work, *in* 'Proceedings of the International Conference of the Learning Sciences', pp. 125–130. 45

Gwizdka, J. (2008), 'Revisiting search task difficulty: Behavioral and individual difference measures', *Proceedings of the American Society for Information Science and Technology* **45**(1), 1–12. 69, 51

Halder, S., Roy, A. and Chakraborty, P. (2010), 'The influence of personality traits on information seeking behaviour of students', *Malaysian Journal of Library and Information Science* **15**(1), 41–53. 3

Hallmark, J. (2001), 'Information-seeking behavior of academic meteorologists and the role of information specialists', *Science & Technology Libraries* **21**(1-2), 53–64. 117

Hartmann, B., Doorley, S. and Klemmer, S. (2008), 'Hacking, mashing, gluing: understanding opportunistic design', *IEEE Pervasive Computing* **7**(3), 46–54. 121, 143, 185, 76, 92

Haynes, R., McKibbon, K., Walker, C., Ryan, N., Fitzgerald, D. and Ramsden, M. (1990), 'Online access to medline in clinical settings', *Annals of Internal medicine* **112**(1), 78. 32, 27

Heider, F. (1944), 'Social perception and phenomenal causality.', *Psychological review* **51**(6), 358. 26, 38, 20, 32

Heinstrm, J. (2003), *Fast surfers, broad scanners and deep divers as users of information technology Relating information preferences to personality traits*, Vol. 40, Wiley Subscription Services, Inc., A Wiley Company.
URL: *http://dx.doi.org/10.1002/meet.1450400131* 3

Heinstrm, J. (2005), 'Fast surfing, broad scanning and deep diving: The influence of personality and study approach on students' information-seeking behavior', *Journal of Documentation* **61**(2), 228–247. 3

Hembrooke, H., Granka, L., Gay, G. and Liddy, E. (2005), 'The effects of expertise and feedback on search term selection and subsequent learning', *Journal of the American Society for Information Science and Technology* **56**(8), 861–871. 11, 3

Hertzum, M. (2002), 'The importance of trust in software engineers' assessment and choice of information sources', *Information and Organization* **12**(1), 1–18. 117

Hertzum, M. and Pejtersen, A. (2000*a*), 'The information-seeking practices of engineers: searching for documents as well as for people', *Information Processing & Management* **36**(5), 761–778. 119

Hertzum, M. and Pejtersen, A. M. (2000*b*), 'The information-seeking practices of engineers: searching for documents as well as for people', *Information Processing & Management* **36**(5), 761–778. 117

Hirsh, S. (2004), 'Domain knowledge and children's search behavior', *Youth information-seeking behavior: Theories, models, and issues* pp. 241–270. 68, 50

Hoc, J. (1990), *Psychology of programming*, Academic Press. 115, 133, 138, 73, 87

Hollan, J., Hutchins, E. and Kirsh, D. (2000), 'Distributed cognition: toward a new foundation for human-computer interaction research', *ACM Transactions on Computer-Human Interaction (TOCHI)* **7**(2), 174–196. 138, 87

Holland, J. and Skinner, B. (1961), *Analysis of behavior.*, McGraw-Hill. 72, 73, 54, 55

Holland, M. P. and Powell, C. K. (1995), 'A longitudinal survey of the information seeking and use habits of some engineers.', *College & Research Libraries* **56**(1), 7–15. 117, 119

Hölscher, C. and Strube, G. (2000), 'Web search behavior of internet experts and newbies', *Computer networks* **33**(1-6), 337–346. 68, 50

Huang, M. and Wang, H. (2004), 'The influence of document presentation order and number of documents judged on users' judgments of relevance', *Journal of the American Society for Information Science and Technology* **55**(11), 970–979. 4, 12, 52, 5, 40

Hunt, J. (1965), 'Traditional personality theory in the light of recent evidence', *American Scientist* **53**, 80. 38, 33

Hutchins, E. (1996), *Cognition in the wild*, The MIT Press. 133, 138, 84, 87

Ingwersen, P. (1996), 'Cognitive perspectives of information retrieval interaction elements of a cognitive ir theory', *Journal of Documentation* **52**, 3–50. 5

Ingwersen, P. and Järvelin, K. (2005*a*), 'Information retrieval in context: Irix', *SIGIR Forum* **39**, 31–39. 94, 231, 173

Ingwersen, P. and Järvelin, K. (2005*b*), *The turn*, Kluwer Academic Pub. 62, 123, 42, 43, 77

Ingwersen, P. and Järvelin, K. (2007), 'On the holistic cognitive theory for information retrieval: Drifting outside the border of the laboratory framework', *Studies in the Theory of Information Retrieval (ICTIR 2007)* pp. 135–147. 12, 14, 15, 61, 62, 94, 96, 5, 7, 8, 40, 42

Jackson, T. (2005), *Motivating sustainable consumption.* 6

Janjic, W., Hummel, O. and Atkinson, C. (2010), More archetypal usage scenarios for software search engines, *in* 'ICSE Workshop on Search-driven Development: Users, Infrastructure, Tools and Evaluation (SUITE '10)', ACM, pp. 21–24. 131, 132, 135, 136, 220, 81, 83, 86, 162

Jenkins, D. P. (2009), *Cognitive work analysis: coping with complexity*, Ashgate Publishing, Ltd. 58

Joachims, T., Granka, L., Pan, B., Hembrooke, H. and Gay, G. (2005), Accurately interpreting clickthrough data as implicit feedback, *in* 'Proceedings of the 28th annual international ACM SIGIR conference on Research and development in information retrieval', SIGIR '05, ACM, New York, NY, USA, pp. 154–161. 4, 11, 12, 52, 3, 5, 40

Joachims, T., Granka, L., Pan, B., Hembrooke, H., Radlinski, F. and Gay, G. (2007), 'Evaluating the accuracy of implicit feedback from clicks and query reformulations in web search', *ACM Transactions on Information Systems (TOIS)* **25**(2), 7–33. 11, 12, 3, 5

Johnson, J. (2000), 'On contexts of information seeking.', *Information Processing and Management* **39**(5), 735–760. 4

Jrvelin, K. and Ingwersen, P. (2004), 'Information seeking research needs extension towards tasks and technology', *Information Research* **10**(1), 361–382. 123, 124, 77, 78

Kapser, C. (2009), Toward an Understanding of Software Code Cloning as a Development Practice, PhD thesis, University of Waterloo. 139, 143, 88, 92

Keane, M., O'Brien, M. and Smyth, B. (2008), 'Are people biased in their use of search engines?', *Communications of the ACM* **51**(2), 49–52. 3, 4, 12, 5

Kellar, M., Watters, C. and Shepherd, M. (2007), 'A field study characterizing web-based information-seeking tasks', *Journal of the American Society for Information Science and Technology* **58**(7), 999–1018. xviii, 174, 175, xiv, 118, 119, 120

Kelley, S. and Mirer, T. (1974), 'The simple act of voting', *The American Political Science Review* **68**(2), 572–591. 63, 45

Kelly, D. and Belkin, N. (2004), Display time as implicit feedback: understanding task effects, *in* 'Proceedings of the 27th annual international ACM SIGIR conference on Research and development in information retrieval', SIGIR '04, ACM, New York, NY, USA, pp. 377–384. 11, 3

Kelly, D. and Cool, C. (2002), The effects of topic familiarity on information search behavior, *in* 'Proceedings of the 2nd ACM/IEEE-CS joint conference on Digital libraries', JCDL '02, ACM, New York, NY, USA, pp. 74–75. 3, 69, 91, 105, 51

Kelly, D. and Teevan, J. (2003), 'Implicit feedback for inferring user preference: a bibliography', *SIGIR Forum* **37**(2), 18–28. 4, 10, 11, 3

Kelly, G. (1992), *The Psychology of Personal Constructs*, Psychology Press. 26, 38, 20, 32

Kim, J. (2006), Task difficulty as a predictor and indicator of web searching interaction, *in* 'CHI '06 Extended Abstracts on Human Factors in Computing Systems', CHI EA '06, ACM, New York, NY, USA, pp. 959–964. 69, 124, 51, 80

KiM, J. (2008*a*), 'Task as a context of information seeking: An investigation of daily life tasks on the web', *Libri* **58**(3), 172–181. 4

Kim, K. (2001*a*), 'Implications of user characteristics in information seeking on the world wide web', *International journal of human-computer interaction* **13**(3), 323–340. 3, 70, 52

Kim, K. (2001*b*), 'Information-seeking on the web: Effects of user and task variables', *Library & Information Science Research* **23**(3), 233–255. 68, 69, 91, 105, 50, 51

Kim, K. (2005), 'Experienced web users' search behavior: Effects of focus and emotion control', *Proc. American Society for Information Science and Technology, Charlotte, USA* . 68, 50

Kim, K. (2008*b*), 'Effects of emotion control and task on web searching behavior', *Information Processing & Management* **44**(1), 373–385. 69, 51

Kim, K. and Allen, B. (2002), 'Cognitive and task influences on web searching behavior', *Journal of the American Society for Information Science and Technology* **53**(2), 109–119. 69, 124, 51, 78

Kim, M., Bergman, L., Lau, T. and Notkin, D. (2004), An ethnographic study of copy and paste programming practices in oopl, *in* 'International Symposium on Empirical Software Engineering (ISESE '04)', IEEE Computer Society, pp. 83–92. 12, 138, 5, 87

King, D. and Griffiths, J. (1991), Indicators of the use, usefulness and value of scientific and technical information, *in* '15th International Online Meeting', Learned Information Ltd., pp. 361–377. 120, 131, 74, 82

REFERENCES

Knussen, C., Yule, F., MacKenzie, J. and Wells, M. (2004), 'An analysis of intentions to recycle household waste: The roles of past behaviour, perceived habit, and perceived lack of facilities', *Journal of Environmental Psychology* **24**(2), 237 – 246. 101

Ko, A. J., Myers, B. A., Coblenz, M. J. and Aung, H. H. (2006), 'An exploratory study of how developers seek, relate, and collect relevant information during software maintenance tasks', *Software Engineering, IEEE Transactions on* **32**(12), 971–987. 127, 129, 130

Kralisch, A. and Berendt, B. (2004), Cultural determinants of search behaviour on websites, *in* 'Proc. of the 6th IWIPS', Vancouver, Canada, pp. 61–74. 4, 72, 54

Kraus, S. (1995), 'Attitudes and the prediction of behavior: A meta-analysis of the empirical literature', *Personality and Social Psychology Bulletin* **21**(1), 58–75. xvii, 9, 35, 36, xiii, 2, 29, 30

Kuhlthau, C. (1991), 'Inside the search process: Information seeking from the user's perspective', *Journal of the American Society for Information Science* **42**(5), 361–371. 59

LaPiere, R. (1934), 'Attitudes vs. actions', *Social forces* **13**(2), 230–237. 33, 28

Large, A., Beheshti, J. and Rahman, T. (2002), 'Gender differences in collaborative web searching behavior: an elementary school study', *Information Processing & Management* **38**(3), 427–443. 4, 72, 54

LaToza, T., Venolia, G. and DeLine, R. (2006), Maintaining mental models: a study of developer work habits, *in* '28th international conference on Software engineering (ICSE '06)', ACM, pp. 492–501. 124, 126, 127, 132, 135, 138, 220, 80, 81, 83, 87, 162

Lawrance, J., Bellamy, R., Burnett, M. and Rector, K. (2008), Using information scent to model the dynamic foraging behavior of programmers in maintenance tasks, *in* 'Proceedings of the twenty-sixth annual SIGCHI conference on Human factors in computing systems', ACM, pp. 1323–1332. 138, 148, 87, 98

Lawrance, J., Bogart, C., Burnett, M., Bellamy, R., Rector, K. and Fleming, S. (2010), 'How programmers debug, revisited: An information foraging theory perspective'. 148, 98

Lazonder, A., Biemans, H. and Wopereis, I. (2000), 'Differences between novice and experienced users in searching information on the world wide web', *Journal of the American Society for Information Science* **51**(6), 576–581. 3

Leckie, G. J. (2005), *Theories of information behavior*, Information Today Inc, theories of information behavior General Model of the Information Seeking of Professionals, p. 431. 117, 118

Leckie, G., Pettigrew, K. and Sylvain, C. (1996), 'Modeling the information seeking of professionals: a general model derived from research on engineers, health care professionals, and lawyers', *The Library Quarterly* pp. 161–193. 117, 118

Lee, H., Belkin, N. and Krovitz, B. (2006), 'Rutgers information retrieval evaluation project on ir performance on different precision levels.', *Journal of the Korean Society for Information Management* **23**(2), 97–111. 31, 25

Lee, T., Park, Y. and Park, Y. (2008), 'A time-based approach to effective recommender systems using implicit feedback', *Expert systems with applications* **34**(4), 3055–3062. 11, 3

Lesk, M. and Salton, G. (1968), 'Relevance assessments and retrieval system evaluation', *Information storage and retrieval* **4**(4), 343–359. 25, 29, 30, 65, 19, 23, 47

Li, N. and Kirkup, G. (2007), 'Gender and cultural differences in internet use: A study of china and the uk', *Computers & Education* **48**(2), 301–317. 4, 72, 54

Li, Y. and Belkin, N. (2008), 'A faceted approach to conceptualizing tasks in information seeking', *Information Processing & Management* **44**(6), 1822–1837. 69, 51

Liu, W., Chen, C., Lakshminarayanan, V. and Perry, D. (2005), A design for evidence-based software architecture research, *in* 'Workshop on REBSE'2005'. 120, 74

MacGregor, J., Ormerod, T. and Chronicle, E. (2001), 'Information processing and insight: a process model of performance on the nine dot and related problems.', *Journal of experimental psychology. Learning, memory, and cognition* **27**(1), 176–201. 69, 51

Manning, C., Raghavan, P. and Schütze, H. (2008), *Introduction to Information Retrieval*. 82, 64

Mantonakis, A., Rodero, P., Lesschaeve, I. and Hastie, R. (2009), 'Order in choice', *Psychological Science* **20**(11), 1309. 4, 52, 40

Marchionini, G. (2006), 'Exploratory search: from finding to understanding', *Communications of the ACM* **49**(4), 41–46. 3, 4

Marchionini, G., Lin, X. and Dwiggins, S. (1990), Effects of search and subject expertise on information seeking in a hypertext environment, *in* 'Proceedings of ASIS (Toronto)', pp. 129–142. 72, 54

Markey, P., Markey, C. and Tinsley, B. (2004), 'Children's behavioural manifestations of the five-factor model of personality.', *Personality and Social Psychology Bulletin* **30**(4), 423–432. 37, 31

McGuire, W. (1960), 'Cognitive consistency and attitude change.', *The Journal of Abnormal and Social Psychology* **60**(3), 345. 26, 38, 20, 32

Minard, R. (1952), 'Race relationships in the pocahontas coal field', *Journal of Social Issues* **8**(1), 29–44. 34, 28

Mischel, W. (1968), *Personality and assessment*, John Wiley and Sons. 38, 32

Mischel, W. (1969), 'Continuity and change in personality.', *American Psychologist* **24**(11), 1012. 37, 38, 31, 32

Mischel, W. and Peake, P. (1982*a*), 'Beyond deja vu in the search for cross-situational consistency', *Psychological Review* **89**, 730. 34, 28

Mischel, W. and Peake, P. (1982*b*), *Consistency in social behavior*, Lawrence Erlbaum Assoc Inc, chapter In search of consistency: Measure for measure, p. 314. 34, 28

Mischel, W. and Peake, P. (1983), 'Some facets of consistency: Replies to epstein, funder, and bem.', *Psychological review* **90**, 394–434. 34, 28

Monson, T., Hesley, J. and Chernick, L. (1982), 'Specifying when personality traits can and cannot predict behavior: An alternative to abandoning the attempt to predict single-act criteria.', *Journal of Personality and Social Psychology* **43**(2), 385. 39, 34

Murphy, J. (2003), 'Information-seeking habits of environmental scientists: A study of interdisciplinary scientists at the environmental protection agency in research triangle park, north carolina', *Issues in Science and Technology Librarianship* **38**, 1–9. 117

Murphy, J., Hofacker, C. and Mizerski, R. (2006), 'Primacy and recency effects on clicking behavior', *Journal of Computer-Mediated Communication* **11**(2), 522–535. 4, 12, 52, 5, 40

Newell, A. and Rosenbloom, P. (1981), 'Mechanisms of skill acquisition and the law of practice', *Cognitive skills and their acquisition* pp. 1–55. 68, 50

Newell, A. and Simon, H. (1972), *Human problem solving*, Prentice Hall. 42, 34

Nisbet, R. and Norenzaya, A. (2004), *Stevens' Handbook of Experimental Psychology, Sensation and Perception*, Wiley, stevens' handbook of experimental psychology, sensation and perception Culture and Cognition, p. 816. 72, 54

Notani, A. (1998), 'Moderators of perceived behavioral controls predictiveness in the theory of planned behavior: A meta-analysis.', *Journal of Consumer Psychology* **7**, 247–271. 103

Oard, D. and Kim, J. (2001), Modeling information content using observable behavior., *in* '64th Annual Meeting of the American Society for Information Science and Technology', pp. 38–45. 4, 12, 5

Ouellette, J. and Wood, W. (1998), 'Habit and intention in everyday life: The multiple processes by which past behavior predicts future behavior.', *Psychological Bulletin* **124**(1), 54–74. 14, 101, 114, 123, 176, 7, 72, 77, 121

Oulasvirta, A., Hukkinen, J. and Schwartz, B. (2009), 'When more is less: the paradox of choice in search engine use', pp. 516–523. 4

Palmer, J. (1991), 'Scientists and information: I. using cluster analysis to identify information style', *Journal of Documentation* **47**(2), 105–129. 117

Paunonen, S. (2003), 'Big five factors of personality and replicated predictions of behavior.', *Journal of Personality and Social Psychology* **84**(2), 411. 9, 36, 2, 30

Philip, K., Umarji, M., Agarwala, M., Sim, S., Gallardo-Valencia, R., Lopes, C. and Ratanotayanon, S. (2012), Software reuse through methodical component reuse and amethodical snippet remixing, *in* 'Proceedings of the ACM 2012 conference on Computer Supported Cooperative Work', ACM, pp. 1361–1370. 138, 143, 87, 92

Pinelli, T. E. (1991), 'The information-seeking habits and practices of engineers', *Science & Technology Libraries* **11**(3), 5–25. 117

Piorkowski, D., Fleming, S., Scaffidi, C., Bogart, C., Burnett, M., John, B., Bellamy, R. and Swart, C. (2012*a*), Reactive information foraging: an empirical investigation of theory-based recommender systems for programmers, *in* 'Proceedings of the 2012 ACM annual conference on Human Factors in Computing Systems', ACM, pp. 1471–1480.
URL: *http://doi.acm.org/10.1145/2208516.2208608*

Piorkowski, D., Fleming, S., Scaffidi, C., Bogart, C., Burnett, M., John, B., Bellamy, R. and Swart, C. (2012*b*), Reactive information foraging: an empirical investigation of theory-based recommender systems for programmers, *in* 'Proceedings of the 2012 ACM annual conference on Human Factors in Computing Systems', ACM, pp. 1471–1480. 148, 98

Pirolli, P. and Card, S. (1999), 'Information foraging.', *Psychological review* **106**(4), 643. 8, 55, 148, 98

Poblete, B. and Baeza-Yates, R. (2008), 'Query-sets: using implicit feedback and query patterns to organize web documents', pp. 41–50. 11, 3

Purgailis Parker, L. and Johnson, R. (1990), 'Does order of presentation affect users' judgment of documents?', *Journal of the American Society for Information Science* **41**(7), 493–494. 3, 4, 12, 5

Rabbitt, P. and Banerji, N. (1989), 'How does very prolonged practice improve decision speed?', *Journal of Experimental Psychology: General* **118**(4), 338. 68, 50

Rees, A. and Schultz, D. (1967), *A field experimental approach to the study of relevance assessments in relation to document searching. Final Report to the National Science Foundation.*, Vol. 1, Clearinghouse for Federal Scientific and Technical Information. 28, 22

Resnick, A. (1961), 'Relative effectiveness of docunient, titles and abstracts for determining relevance of documents.', *Science* pp. 1004–1005. 28, 22

Resnick, A. and Savage, T. (1960), 'A re-evaluation of machine-generated abstracts.', *Human Factors* pp. 141–146. 28, 22

Resnick, A. and Savage, T. (1964), 'The consistency of human judgments of relevance.', *American Documentation* **15**(2), 93–95. 28, 33, 22, 27

Reynolds, T. J. and Gutman, J. (1988), 'Laddering theory, method, analysis, and interpretation', *Journal of advertising research* **28**(1), 11–31. 6

Rijsbergen, C. J. (2004), *The geometry of information retrieval*, Cambridge University Press. 72, 54

Rivis, A. and Sheeran, P. (2003), 'Descriptive norms as an additional predictor in the theory of planned behaviour: A meta-analysis.', *Current Psychology* **22**, 218–233. 103

Robertson, S. (1999), *Types of thinking*, Psychology Press. 43, 36

Rosenbaum, H., Davenport, E., Lievrouw, L. A. and Day, R. E. (2003), 'The death of the user. sponsored by sig hfis, crit', *Proceedings of the American Society for Information Science and Technology* **40**(1), 429–430.
URL: *http://dx.doi.org/10.1002/meet.1450400153* 43, 44

REFERENCES

Roy, M. and Chi, M. (2003), 'Gender differences in patterns of searching the web', *Journal of Educational Computing Research* **29**(3), 335–348. 4, 72, 54

Rushton, J., Brainerd, C. and Pressley, M. (1983), 'Behavioral development and construct validity: The principle of aggregation', *Psychological Bulletin* **94**(1), 18–38. 97

Sandhusen, R. (2008), *Marketing*, Barrons Educational Series. 73, 55

Saracevic, T. (1992), 'Encyclopedia of library and information sciences'. 117

Saracevic, T. (1996), 'Modeling interaction in information retrieval (ir): A review and proposal.', *Proceedings of the American Society for Information Science* **33**, 3–9. 12, 14, 5, 7

Saracevic, T. (1997*a*), 'The stratified model of information retrieval interaction: Extension and applications', *Proceedings of the American Society for Information Science* **34**, 313–327. 12, 14, 52, 53, 5, 7, 38, 41

Saracevic, T. (1997*b*), 'Users lost. reflections on the past, future, and limits of information science.', *SIGIR Forum* **31**(2), 16–27. 5

Saracevic, T. (2007), 'Relevance: A review of the literature and a framework for thinking on the notion in information science. part iii: Behavior and effects of relevance', *Journal of the American Society for Information Science and Technology* **58**(13), 2126–2144. 4, 9, 27, 2, 21

Saracevic, T., Kantor, P., Chamis, A. and Trivison, D. (1988), 'A study of information seeking and retrieving: I. background and methodology', *Journal of the American Society for Information Science* **39**, 176–195. 31, 25

Saracevic, T., Kantor, P., Chamis, A. and Trivison, D. (1997), 'A study of information seeking and retrieving: 1. background and methodology', *Readings in Information Retrieval. San Francisco: Morgan Kaufmann* pp. 175–190. 3, 231, 173

Savolainen, R. (1995), 'Everyday life information seeking: approaching information seeking in the context of 'way of life'', *Library & Information Science Research* **17**(3), 259–294. 12, 14, 47, 48, 49, 5, 7, 38, 39

Sawyer, S. (2004), 'Software development teams', *Communications of the ACM* **47**(12), 95–99. 136, 87

Schmidt, C. F., Sridharan, N. S. and Goodson, J. L. (1978), 'The plan recognition problem: An intersection of psychology and artificial intelligence', *Artif. Intell.* **11**(1-2), 45–83. 44

Schneider, W. and Chein, J. (2003), 'Controlled & automatic processing: behavior, theory, and biological mechanisms', *Cognitive Science* **27**(3), 525–559. 39, 34

Schulze, R. and Wittmann, W. (2003), *Meta-Analysis: New Developments and Applications in Medical and Social Sciences*, Hogrefe Verlag GmbH + Company, chapter A meta-analysis of the theory of reasoned action and the theory of planned behavior: The principle of compatibility and multidimensionality of beliefs as moderators., pp. 219–250. 14, 103, 7

Scott, J. (2000), 'Rational choice theory', *Understanding contemporary society: Theories of the present* pp. 126–138. 6

Seaman, C. (2002), The information gathering strategies of software maintainers, *in* 'International Conference on Software Maintenance (ICSM'02)', IEEE Computer Society, pp. 141–150. 120, 131, 134, 74, 82, 84

Seibel, R. (1963), 'Discrimination reaction time for a 1,023-alternative task.', *Journal of Experimental Psychology* **66**(3), 215. 68, 50

Seligman, M. and Miller, S. (1979), *Choice and perceived control.*, Lawrence Erlbaum Associates, chapter The psychology of power: Concluding comments. 63, 45

Sharif, K. Y., Komputer, F. S. and Maklumat, T. (2011), 'Open source programmers information seeking during software maintenance', *Journal of Computer Science* **7**(7), 1060–1071. 127, 128

Shaw, W., Wood, J., Wood, R. and Tibbo, H. (1991), 'The cystic fibrosis database: Content and research opportunities', *Library and information science research* **13**(4), 347–366. 30, 24

Sihvonen, A. and Vakkari, P. (2004), 'Subject knowledge improves interactive query expansion assisted by a thesaurus', *Journal of Documentation* **60**(6), 673–690. 3

Sim, S., Clarke, C. and Holt, R. (1998), Archetypal source code searches: A survey of software developers and maintainers, *in* '6th International Workshop on Program Comprehension (IWPC '98)', IEEE Computer Society, pp. 180–187. 132, 134, 83, 85

Sim, S., Umarji, M., Ratanotayanon, S. and Lopes, C. (2011), 'How well do search engines support code retrieval on the web?', *ACM Transactions on Software Engineering and Methodology (TOSEM)* **21**(1), 4. 139, 88

Simon, H. (1973), 'The structure of ill structured problems', *Artificial Intelligence* **4**(3-4), 181–201. 69, 51

Singer, J., Lethbridge, T., Vinson, N. and Anquetil, N. (1997), An examination of software engineering work practices, *in* J. H. Johnson, ed., 'Conference of the Centre for Advanced Studies on Collaborative research (CASCON '97)', IBM Press, pp. 174–188. 120, 121, 131, 134, 74, 75, 82, 84

Skinner, B. (1938), *The Behavior of Organisms: An Experimental Analysis.*, B.F. Skinner Foundation. 72, 54

Snyder, M. (1974), 'Self-monitoring of expressive behavior.', *Journal of personality and social psychology* **30**(4), 526. 39, 34

Sormunen, E. (2002), Liberal relevance criteria of trec: Counting on negligible documents?, *in* '25st Annual International Conference on Research and Development in Information Retrieval of the Special Interest Group on Information Retrieval', Association for Computing Machinery, pp. 324–330. 32, 26

Spink, A. and Cole, C. (2005), *New directions in cognitive information retrieval*, Kluwer Academic Pub. 12, 14, 40, 5, 7, 34

Spink, A. and Heinström, J. (2011), *New Directions in Information Behaviour*, Emerald Group Publishing Limited. 12, 14, 40, 5, 7, 34

Spink, A. and Losee, R. (1996), 'Feedback in information retrieval', *Annual Review of Information Science and Technology* **31**, 33–78. 11, 3

Suchman, L. (2007), *Human-Machine Reconfigurations: Plans and Situated Actions*, Learning in Doing: Social, Cognitive and Computational Perspectives, Cambridge University Press.
URL: *http://books.google.co.uk/books?id=VwKMDV-Gv1MC* 44, 102

Suchman, L. A. (1987), *Plans and situated actions - the problem of human-machine communication*, Learning in doing: social,cognitive,and computational perspectives, Cambridge University Press. 44

Sutcliffe, A. and Ennis, M. (1998), 'Towards a cognitive theory of information retrieval', *Interacting with computers* **10**(3), 321–351. 59

Sutcliffe, A. G., Ennis, M. and Watkinson, S. J. (2000), 'Empirical studies of end-user information searching', *Journal of the American Society for Information Science* **51**(13), 1211–1231.
URL: *http://dx.doi.org/10.1002/1097-4571(2000)9999:9999¡::AID-ASI1033¿3.0.CO;2-5* 59

Taylor, R. S. (1962), 'The process of asking questions', *American Documentation* **13**(4), 391–396.
URL: *http://dx.doi.org/10.1002/asi.5090130405* 45

Teevan, J., Dumais, S. and Horvitz, E. (2010), 'Potential for personalization', *ACM Transactions on Computer-Human Interaction (TOCHI)* **17**(1). 2, 1

Tenopir, C. and King, D. W. (2004), *Communication patterns of engineers*, Wiley-IEEE Press. 117, 119

Trafimow, D. and Sheeran, P. (1998), 'Some tests of the distinction between cognitive and affective beliefs', *Journal of Experimental Social Psychology* **34**(4), 378–397. 101

Tu, Y., Shih, M. and Tsai, C. (2008), 'Eighth graders' web searching strategies and outcomes: The role of task types, web experiences and epistemological beliefs', *Computers & Education* **51**(3), 1142–1153. 3, 4, 69, 70, 71, 51, 52, 53

Turpin, A. and Scholer, F. (2006), User performance versus precision measures for simple search tasks, *in* 'Proceedings of the 29th annual international ACM SIGIR conference on Research and development in information retrieval', SIGIR '06, ACM, New York, NY, USA, pp. 11–18. 69, 51

Tversky, A. and Kahneman, D. (1992), 'Advances in prospect theory: Cumulative representation of uncertainty', *Journal of Risk and uncertainty* **5**(4), 297–323. 39, 33

Umarji, M., Sim, S. and Lopes, C. (2008), Archetypal internet-scale source code searching, *in* B. Russo, E. Damiani, S. Hissam, B. Lundell and G. Succi, eds, 'Open Source Development, Communities and Quality', Vol. 275 of *IFIP The International Federation for Information Processing*, Springer US, pp. 257–263. xvii, 129, 132, 134, 135, 138, xiii, 81, 83, 85, 87

Vakkari, P. (2003), 'Task-based information searching', *Annual review of information science and technology* **37**(1), 413–464. 59, 60

Vakkari, P. (2008), 'Trends and approaches in information behaviour research', *Information Research* **13**(4), 13–4. 59

Vakkari, P. and Sormunen, E. (2004), 'The influence of relevance levels on the effectiveness of interactive information retrieval', *Journal of the American Society for Information Science and Technology* **55**(11), 963–969. 32, 41, 26

Vicente, K. (1999), *Cognitive work analysis*, Lawrence Erlbaum. 56, 43

von Mayrhauser, A. (1995), 'Program comprehension during software maintenance and evolution', *Computer* pp. 44–55. 115, 133, 73, 84

Voorhees, E. (1998), Variations in relevance judgments and the measurement of retrieval effectiveness, *in* 'Proceedings of the 21st annual international ACM SIGIR conference on Research and development in information retrieval', ACM Press, pp. 315–323. 32, 26

Voorhees, E. (2000), 'Variations in relevance judgments and the measurement of retrieval effectiveness', *Information Processing and Management* **36**(5), 697–716. 32, 26

Voorhees, E. (2001), Evaluation by highly relevant documents, *in* '24th Annual International ACM SIGIR Conference on Research and Development in Information Retrieval', ACM Press, pp. 74–82. 32, 26

Voorhees, E., Harman, D., of Standards, N. I. and (U.S.), T. (2005), *TREC: experiment and evaluation in information retrieval*, Digital libraries and electronic publishing, MIT Press. 32, 26

Voss, J. and Post, T. (1988), *The Nature of Expertise*, Lawrence Erlbaum Assoc Inc, the nature of expertise On the solving of ill-structured problems. The nature of expertise., p. 434. 69, 51

Weinberg, G. (1998), *The psychology of computer programming*, Dorset House. 115, 133, 138, 73, 84, 87

Weiser, E. (2000), 'Gender differences in internet use patterns and internet application preferences: A two-sample comparison', *Cyberpsychology and behavior* **3**(2), 167–178. 72, 54

Wersig, G. (1979), 'The problematic situation as a basic concept of information science in the framework of social sciences: A reply to belkin, n.j.', *Theoretical problems of informatics: New trends in informatics and its terminology* . 45

White, H. D. and McCain, K. W. (1998), 'Visualizing a discipline: An author co-citation analysis of information science, 1972-1995', *Journal of the American society for information science* **49**(4), 327–355. 5, 40

White, R. and Drucker, S. (2007), Investigating behavioral variability in web search, *in* 'Proceedings of the 16th international conference on World Wide Web', WWW '07, ACM, New York, NY, USA, pp. 21–30. 68, 72, 50, 54

White, R., Dumais, S. and Teevan, J. (2009), Characterizing the influence of domain expertise on web search behavior, *in* R. Baeza-Yates, P. Boldi, B. Ribeiro-Neto and B. Cambazoglu, eds, 'Second ACM International Conference on Web Search and Data Mining (WSDM '09)', ACM, pp. 132–141. 3, 68, 50

REFERENCES

White, R. and Horvitz, E. (2009), 'Experiences with web search on medical concerns and self diagnosis.', *AMIA Symposium* **2009**, 696–700. 68, 50

White, R. and Kelly, D. (2006), A study on the effects of personalization and task information on implicit feedback performance, *in* 'Proceedings of the 15th ACM international conference on Information and knowledge management', CIKM '06, ACM, New York, NY, USA, pp. 297–306. 11, 3

White, R. and Morris, D. (2007), 'Investigating the querying and browsing behavior of advanced search engine users', pp. 255–262. 3

White, R. and Roth, R. (2009), *Exploratory search*, Morgan and Claypool. 4, 70, 71, 52, 53

Wicker, A. (1969), 'Attitudes versus actions: The relationship of verbal and overt behavioral responses to attitude objects', *Journal of social issues* **25**(4), 41–78. 9, 33, 35, 37, 38, 63, 2, 27, 29, 31, 32, 45

Wildemuth, B. (2004), 'The effects of domain knowledge on search tactic formulation', *Journal of the American Society for Information Science and Technology* **55**(3), 246–258. 68, 50

Wilson, T. (1999), 'Models in information behaviour research', *Journal of documentation* **55**(3), 249–270. 12, 14, 49, 50, 51, 5, 7, 36, 37

Wilson, T. (2000), 'Human information behavior', *Informing science* **3**(2), 49–56. 3, 4, 15, 8

Wilson, T. (2008), 'The information user past present and future', *Journal of Information Science* **34**(4), 457–464. 5

Wilson, T., Dunn, D., Bybee, J., Hyman, D. and Rotondo, J. (1984), 'Effects of analyzing reasons on attitude-behavior consistency.', *Journal of Personality and Social Psychology* **47**(1), 5. 116

Witkin, H. and Berry, J. (1975), 'Psychological differentiation in cross-cultural perspective.', *Journal of cross-cultural psychology* . 68, 50

Witkin, H. and Goodenough, D. (1981), *Cognitive styles, essence and origins*, Intl Universities Pr Inc. 68, 50

Witkin, H., Hansson, S. B. and Rydèn, O. (1980), *Psychological differentiation: studies of development*, Wiley. 68, 50

Wood, F., Ford, N., Miller, D., Sobczyk, G. and Duffin, R. (1996), 'Information skills, searching behaviour and cognitive styles for student-centred learning: a computer-assisted learning approach', *Journal of Information Science* **22**(2), 79. 3

Wood, W., Quinn, J. and Kashy, D. (2002), 'Habits in everyday life: Thought, emotion, and action.', *Journal of Personality and Social Psychology* **83**(6), 1281. xviii, 14, 101, 114, 123, 176, 179, 180, 181, 201, xiv, 7, 72, 77, 121, 122, 124, 125, 143

Xu, S. and Rajlich, V. (2004), Cognitive process during program debugging, *in* 'Cognitive Informatics, 2004. Proceedings of the Third IEEE International Conference on', IEEE, pp. 176–182. 115, 133, 73, 84

Xu, Y. and Wang, D. (2008), 'Order effect in relevance judgment', *Journal of the American Society for Information Science and Technology* **59**(8), 1264–1275. 3, 4, 12, 5

Ye, Y., Yamamoto, Y. and Nakakoji, K. (2007), A socio-technical framework for supporting programmers, *in* '6th joint meeting of the European software engineering conference and the ACM SIGSOFT symposium on The foundations of software engineering (ESEC-FSE '07)', ACM, pp. 351–360. 115, 73

Yeh, R., Paepcke, A. and Klemmer, S. (2008), Iterative design and evaluation of an event architecture for pen-and-paper interfaces, *in* '21st annual ACM symposium on User interface software and technology (UIST '08)', ACM, pp. 111–120. 138, 87

Yitzhaki, M. and Hammershlag, G. (2004), 'Accessibility and use of information sources among computer scientists and software engineers in israel: Academy versus industry', *Journal of the American Society for Information Science and Technology* **55**(9), 832–842. 117

Zhang, B.-T. and Seo, Y. (2001), 'Personalized web-document filtering using reinforcement learning', *Applied Artificial Intelligence* **15**, 665–685. 12, 5

Zhang, X., Anghelescu, H. and Yuan, X. (2005), 'Domain knowledge, search behaviour, and search effectiveness of engineering and science students: an exploratory study', *Information Research* **10**(2), 10–2. 3, 68, 50

Zipf, G. (1949), *Human behavior and the principle of least effort*, Martino Fine Books. 55, 136, 87

Zunde, P. and Dexter, M. (1969), 'Indexing consistency and quality', *American Documentation* **20**(3), 259–267. xi, 29, 23